



深層学習を用いたロボットによる 高速汎用物体操作

筑波大学システム情報系
知能機能工学域
准教授 境野翔

社会的背景

未だにロボット化されていない肉体労働は多い



エンジンの組立



きゅうりの箱詰め



倉庫内の陳列



建築



料理



手術

物体の配置・形状・剛性が毎回微妙に異なる

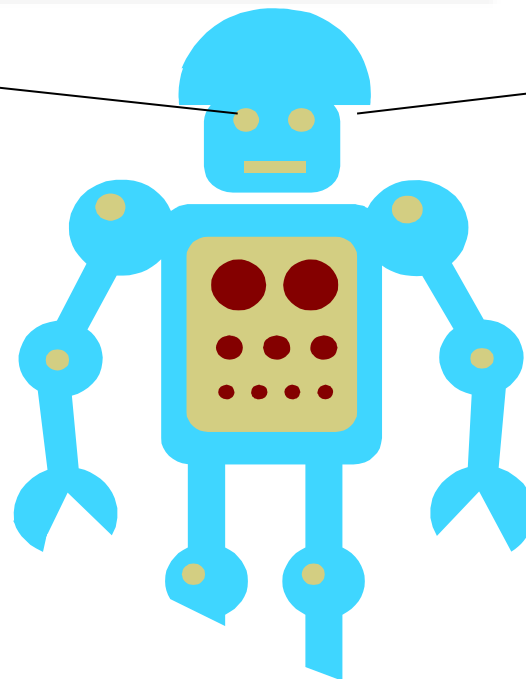
未知物体の操作の必要なタスク

ロボットの現状

深層学習(Deep Learning)では
賢い認識・判断はできるが触れない

画像認識
DLで大幅に向上

音声認識
DLで大幅に向上

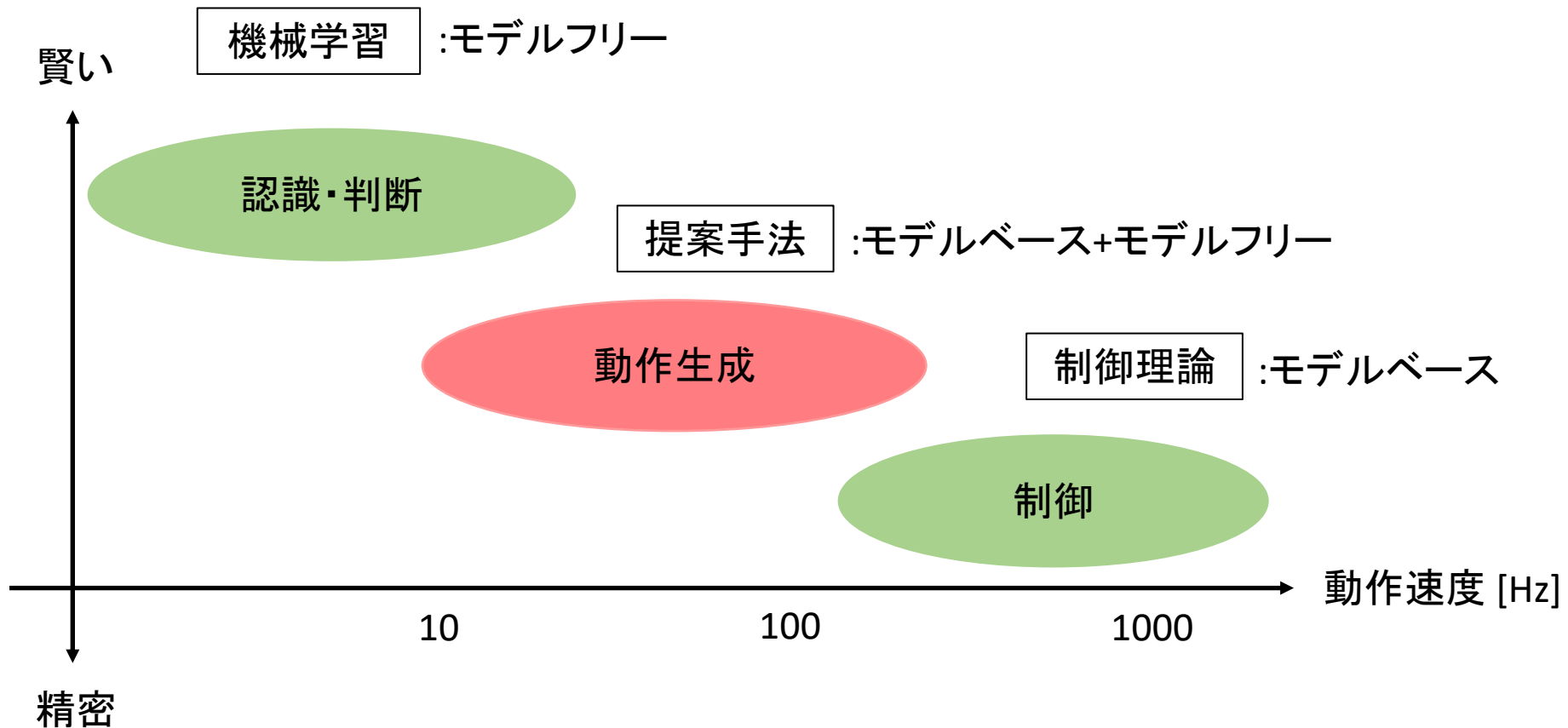


動作計画
接触を含む動作計画は困難

ラスト5cm問題

動作生成

本研究のターゲット

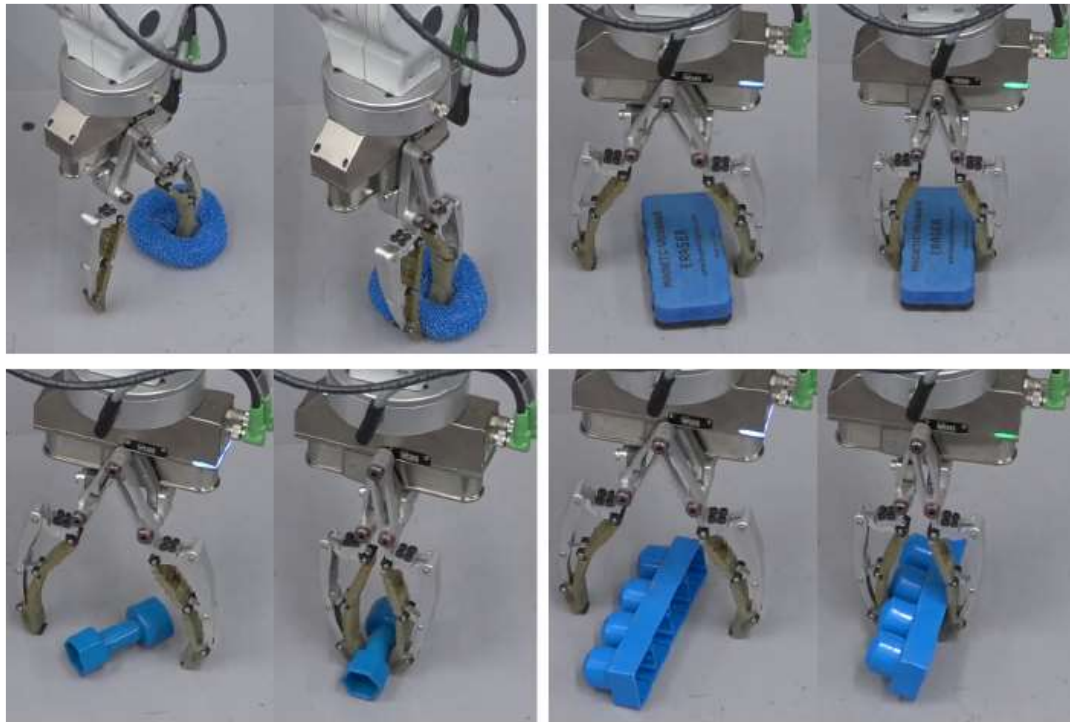


本技術の効果

	従来手法	境野従来技術	新技術
環境への適応能力	低	高	高
動作速度	遅(人間の1/5程度)	速(人間並)	速(人間並)
長周期の動作	可能(数分程度)	困難(10秒程度)	可能(数分程度)
時間方向の汎化性能	無	無	有
AIの設計コスト	高	高	低

従来手法1(強化学習)

強化学習には膨大な試行回数が必要

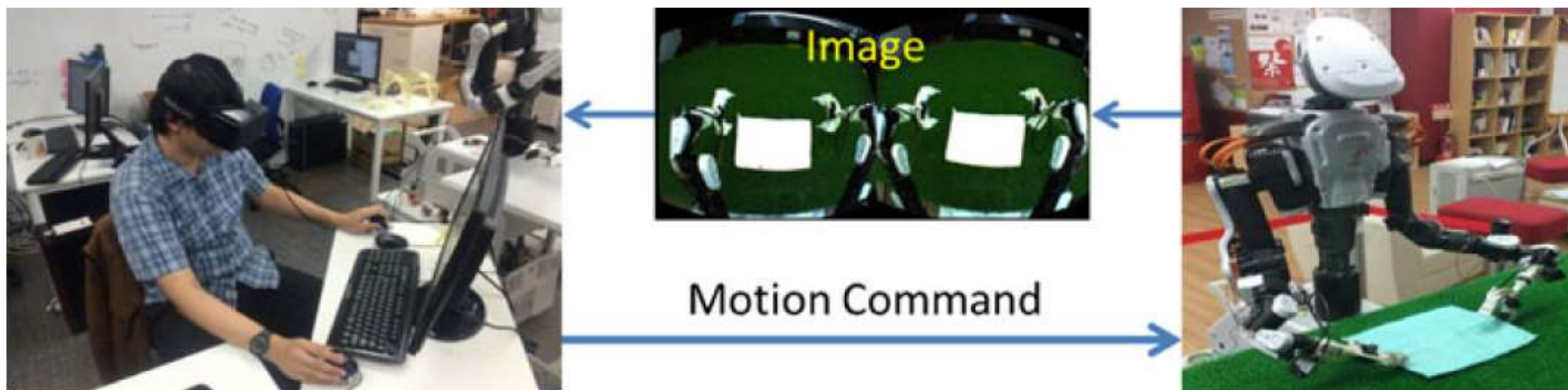


*14台のロボットで80万回試行し強化学習

従来手法2(模倣学習)

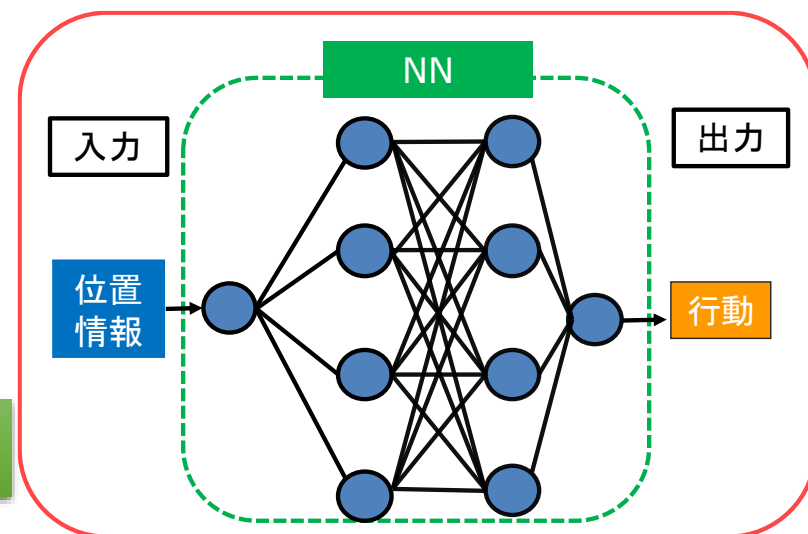
人間の位置指令をデータとして収集

[*]



40回の教示データより
模倣学習(教師あり学習)

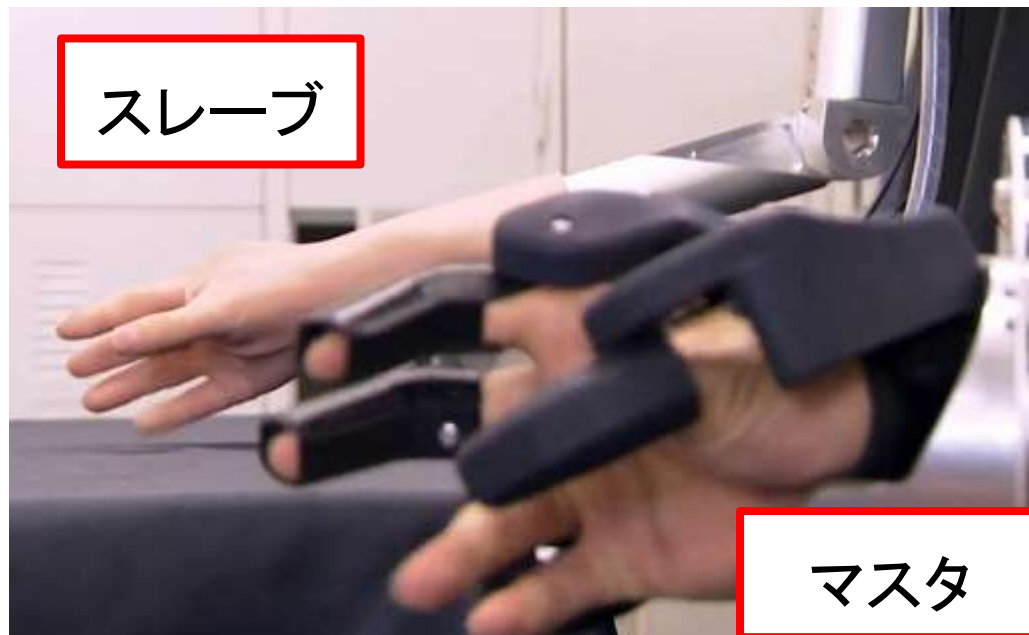
しかし、物体操作成功率8割弱



[*] Pin-Chu Yang, Kazuma Sasaki, Kanata Suzuki, Kei Kase, Shigeki Sugano, Tetsuya Ogata: "Repeatable Folding Task by Humanoid Robot Worker Using Deep Learning," IEEE Robotics and Automation Letters, Vol. 2, No. 2 2017.

バイラテラル制御

バイラテラル制御によるポテトチップスの把持*

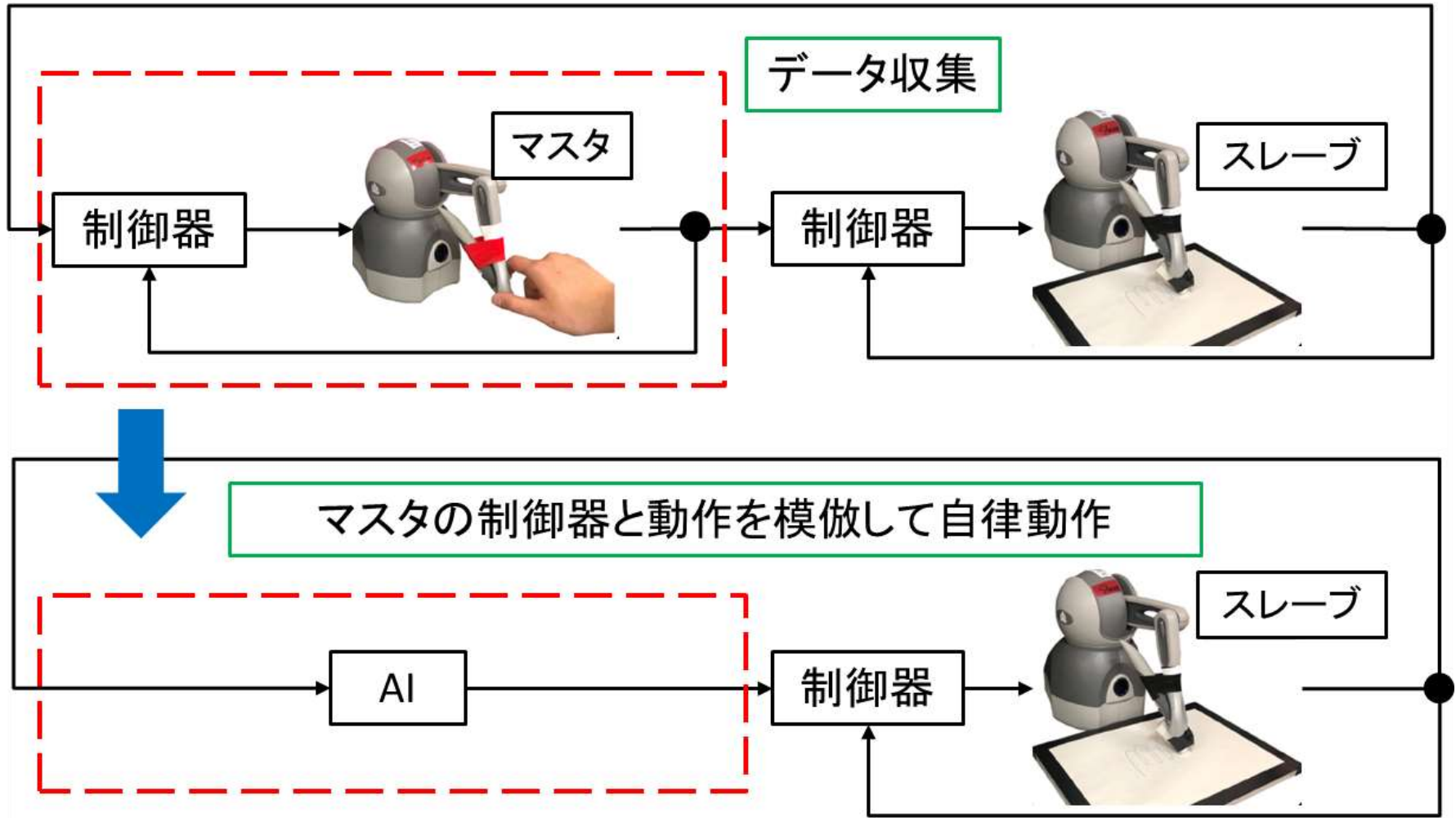


[*] 慶應義塾大学野崎研究室, <http://www.fha.sd.keio.ac.jp/jp/nozaki/nozaki.html>



人間がロボットを遠隔操作すれば柔軟な把持は可能

バイラテラル制御を用いた模倣学習 (境野の過去の発明技術)



バイラテラル制御を用いた模倣学習

バイラテラル制御で教師データを収集

バイラテラル制御

力情報をも利活用

データベース

位置情報

力情報

学習

ニューラルネットワーク

入力

位置情報

力情報

出力

位置指令値

力指令値

位置制御と
力制御を再現可能

バイラテラル制御を用いた模倣学習

- 発明の名称：行動推定装置、行動推定方法、および行動推定プログラム
- 出願番号：特願2018- 78057
- 出願人：埼玉大学
- 発明者：境野翔

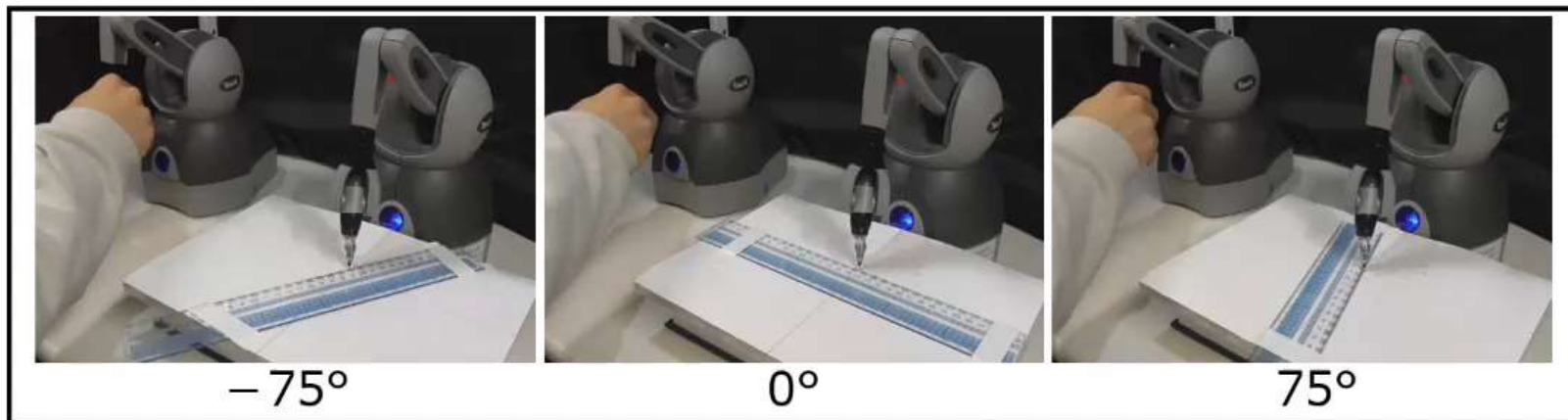
バイラテラル制御を用いた模倣学習

反力呈示する遠隔操作「バイラテラル制御」
を用いて動作を教示

Step1. 複数の異なる定規の傾きで線を引く動作を学習

学習する定規の傾き：15°間隔で-75°~75°（計11種類）

例



遠隔操作技術のバイラテラル制御を用いてデータを取得・保存

力加減を教示可能

バイラテラル制御を用いた模倣学習

1次元的教师データから
2次元の動作を獲得

Step2. 学習モデルに従い自律動作を実行

未学習の傾き: -40°



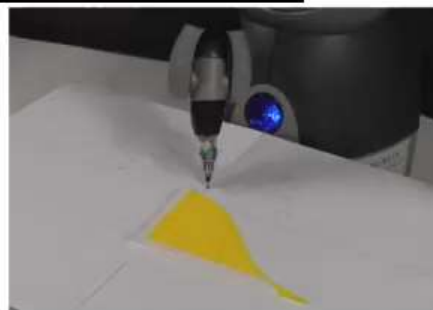
分度器→定規 (未学習)



未学習の傾き: 20°



曲線1 (未学習)



曲線2 (未学習)

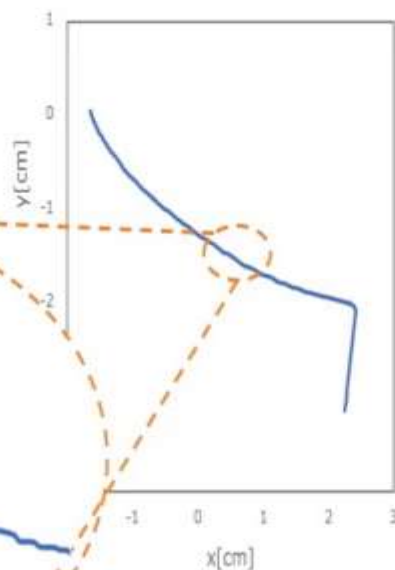


バイラテラル制御を用いた模倣学習

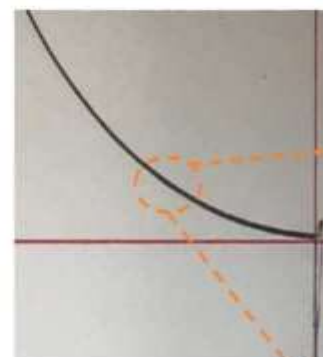


道具と力制御を使うとロボットの位置制御性能を超越可能

Q 拡大表示



ロボットの応答値
(角度センサから取得)

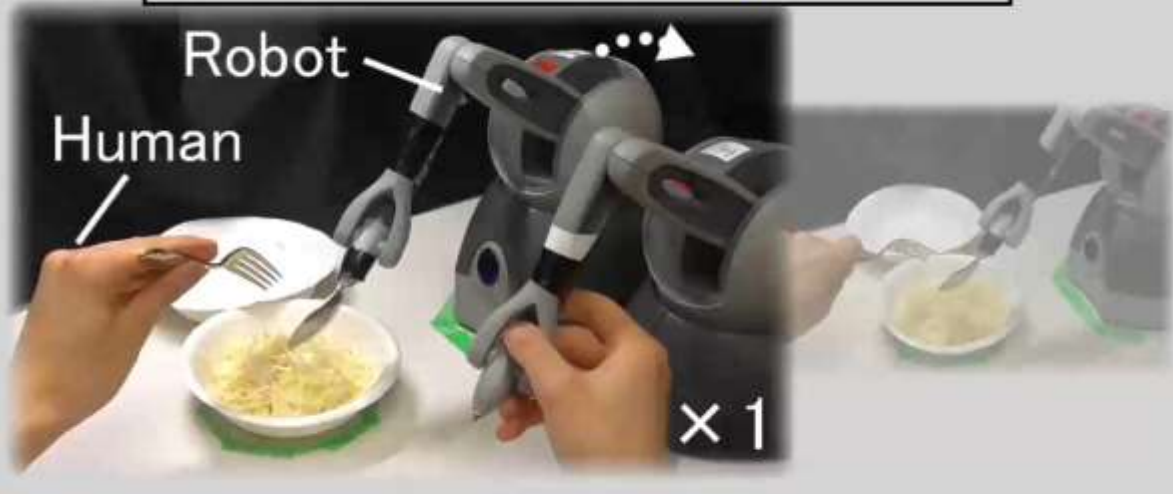


分度器に沿って引かれた線

バイラテラル制御を用いた模倣学習

人間との協調作業

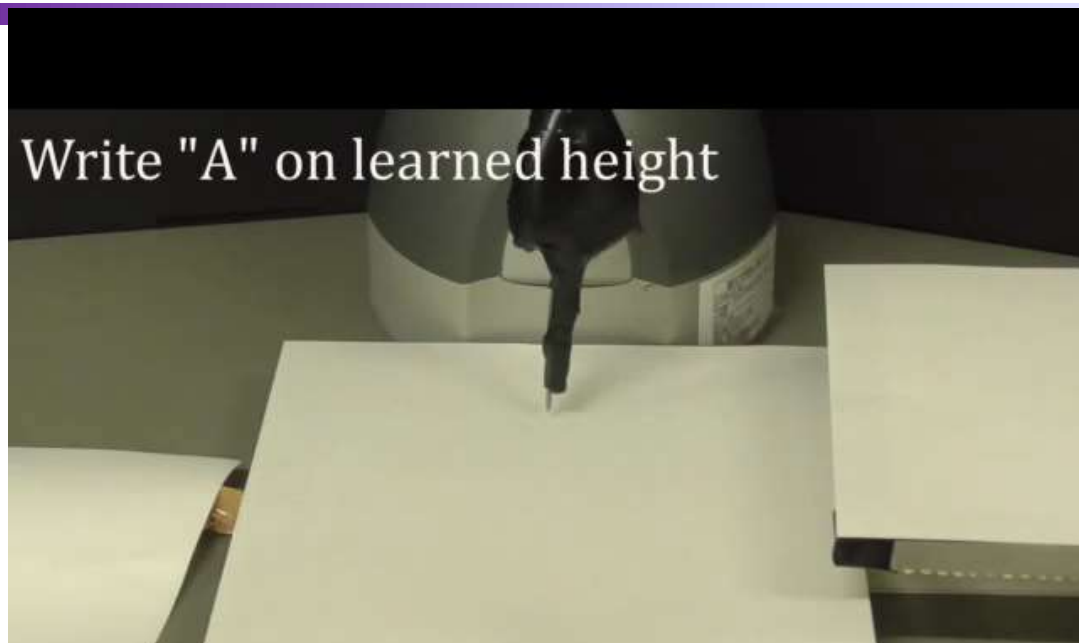
Human-Robot Cooperation



Training..



バイラテラル制御を用いた模倣学習



バイラテラル制御を用いた模倣学習

- 人間並みの速くてなめらかな動作
- 力制御で未知環境へ適応
- 非常に少ない(せいぜい数十回程度)動作教示

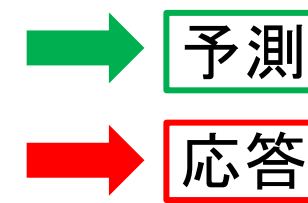
教示動作を変更するだけで汎用的な技能を獲得

ハードウェアの変更は必要なし

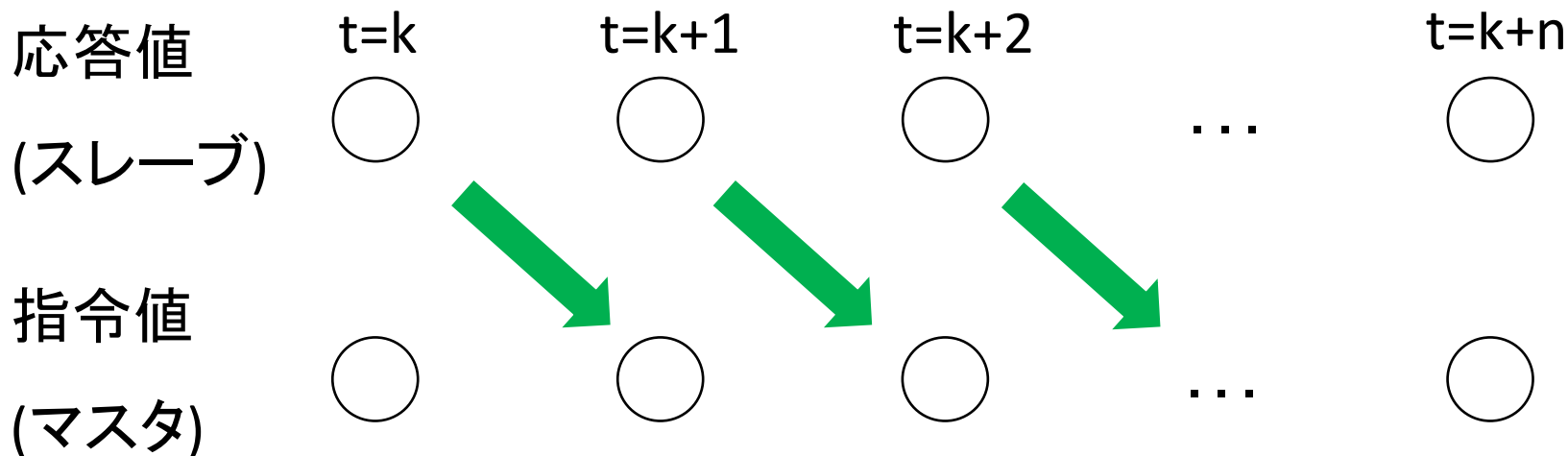
問題点

- 長期間の動作の生成が不安定である
- 空間方向の汎化性能にのみ注力しており
時間方向の汎化性能を考慮していない
- 動作の生成に認識から制御まですべてを設計する必要がある

不安定な長期動作生成

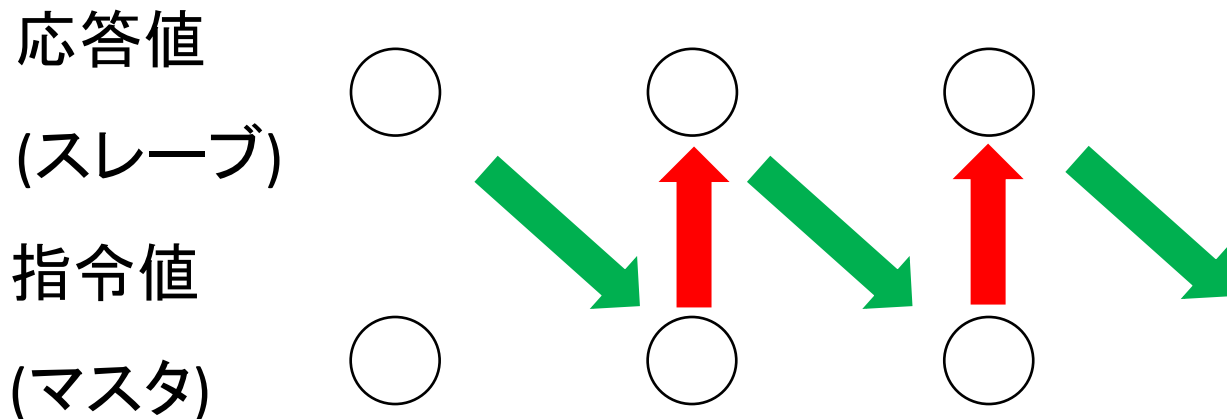


学習時



実行時

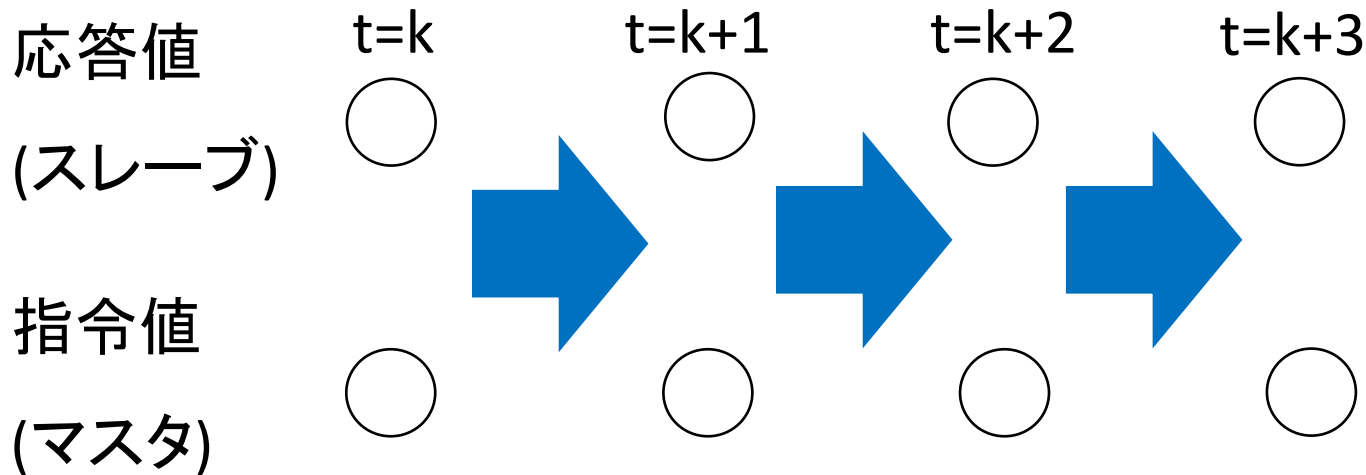
$t=k$ の状態から $t=k+n$ の状態を予測不可能



学習時:開ループ
実行時:閉ループ

自己回帰モデルによる長期動作生成

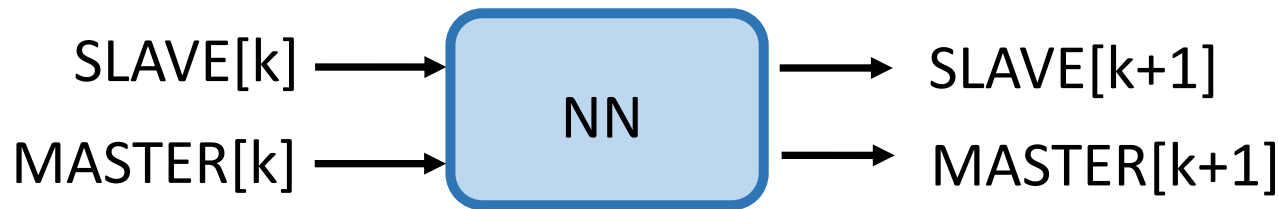
学習時



$$\text{Loss} = \text{Loss}[k+1] + \text{Loss}[k+2] + \text{Loss}[k+3] \dots$$

t=kのスレーブから次ステップのマスタと
スレーブを予測

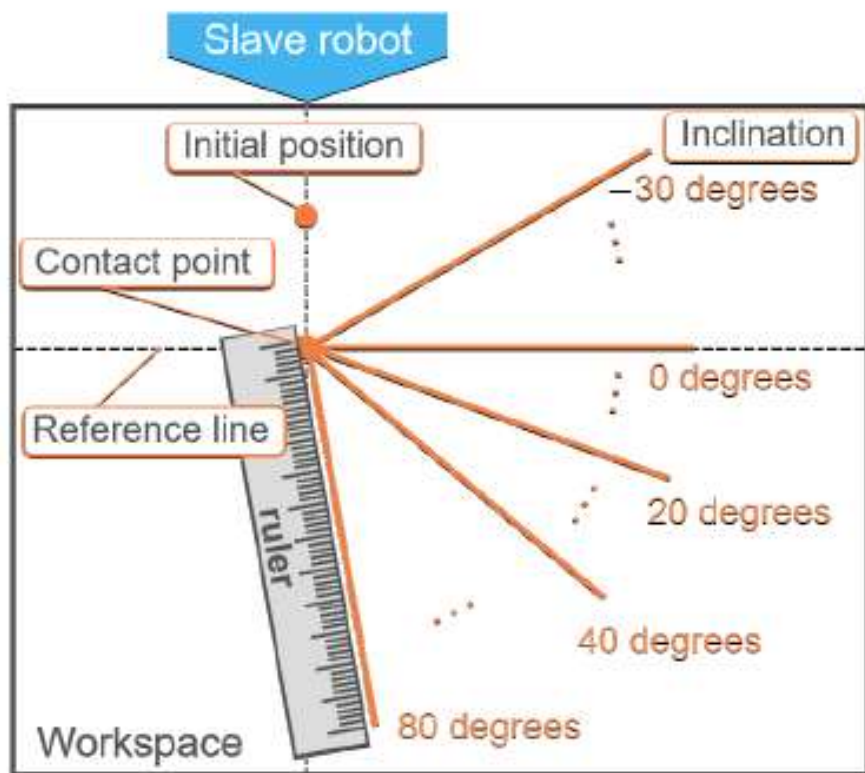
自己回帰モデルによる長期動作生成



- ・スレーブとマスタの将来を同時に予測することで、これらに共通する特徴量を抽出できるため、汎化性能向上
- ・長期の誤差を減らすように学習できるため、ブレが少なく安定な動作を生成可能

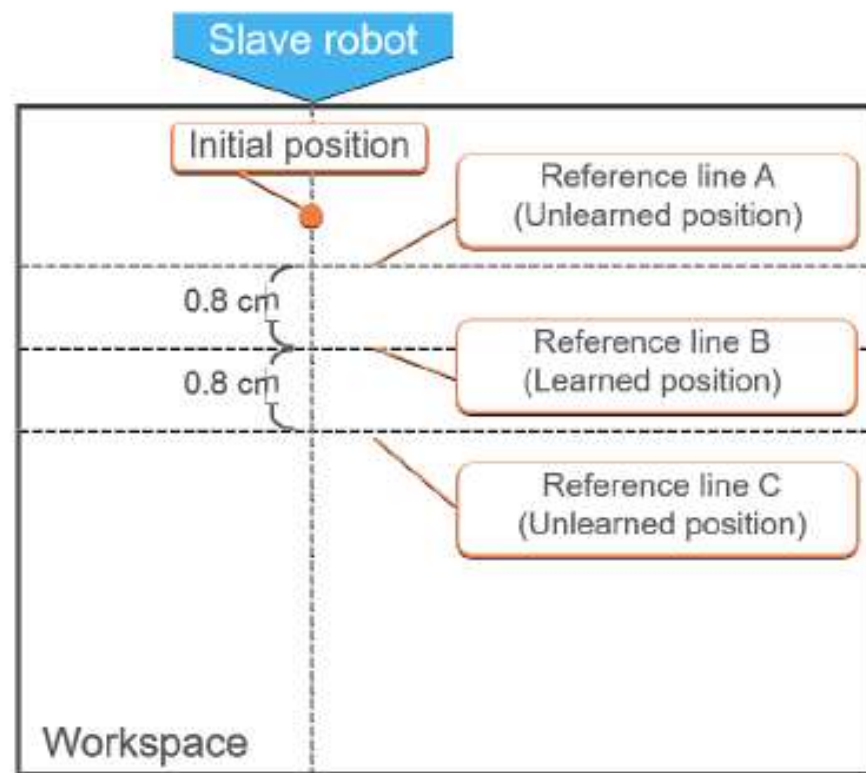
定規を使うタスクによる実証

教師データは傾き 0° , 20° , 40°



(a) 定規の傾き

教師データはBの位置



(b) 初期位置から定規までの距離

定規を使うタスクによる実証

描線成功率

Model	Reference line	Success Rate [%]												Subtotal	Total
		Inclination [deg]													
		-30	-20	-10	0*	10	20*	30	40*	50	60	70	80		
S2S-w/o-AR	A	0	100	100	100	100	100	100	100	100	100	100	100	91.7	88.0 (95/108)
	B**	0	100	100	100	100	100	100	100	100	100	100	100	91.7	
	C	0	0	100	100	100	100	100	100	100	100	100	66.7	81.0	
S2S-AR	A	0	100	100	100	100	100	100	100	100	100	100	100	91.7	93.5 (101/108)
	B**	0	100	100	100	100	100	100	100	100	100	100	100	91.7	
	C	100	100	100	100	100	100	100	100	100	100	100	66.7	97.2	
S2M-w/o-AR	A	0	0	0	66.7	100	100	100	100	100	100	100	66.7	69.4	66.7 (72/108)
	B**	0	0	0	100	100	100	100	100	100	100	66.7	0	63.9	
	C	0	0	0	100	100	100	100	100	100	100	100	0	66.7	
SM2SM-w/o-AR	A	0	0	100	100	100	100	100	100	0	100	100	100	75.0	81.0 (87/108)
	B**	0	0	100	100	100	100	100	100	100	100	100	100	83.3	
	C	0	0	100	100	100	100	100	100	100	100	100	100	83.3	
SM2SM-AR (Proposed method)	A	100	100	100	100	100	100	100	100	100	100	100	100	100	100 (108/108)
	B**	100	100	100	100	100	100	100	100	100	100	100	100	100	
	C	100	100	100	100	100	100	100	100	100	100	100	100	100	

*: Learned inclination of the ruler

** : Learned position of the ruler

外挿時も成功率100%

S2S: 従来手法(スレーブからスレーブを予測)

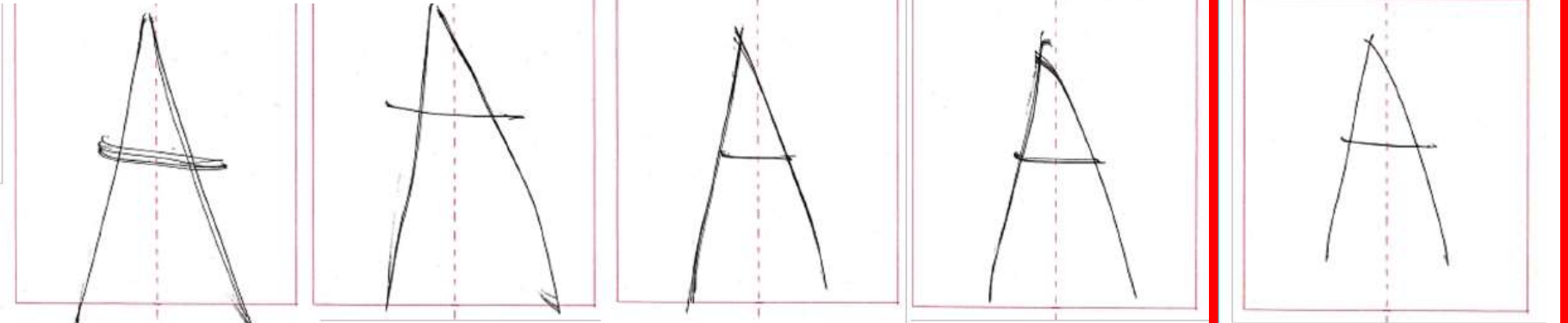
S2M: 境野の従来手法(スレーブからマスタを予測)

SM2SM: 提案手法(スレーブとマスタからスレーブとマスタを予測)

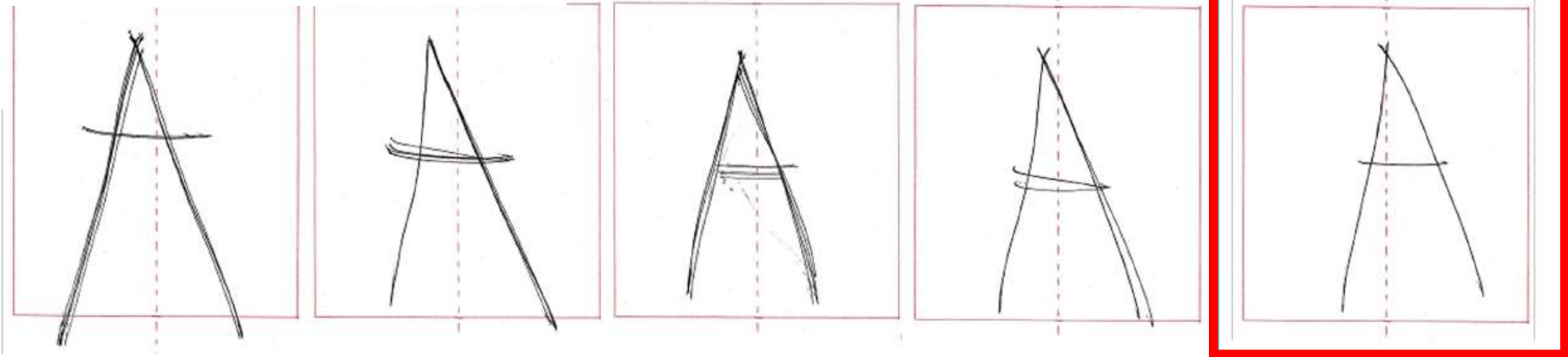
AR: 自己回帰モデル(長期的な誤差を予測)

字を書くタスクによる実証

教師データ: 40文字



教師データ: 4文字



従来手法1

従来手法2

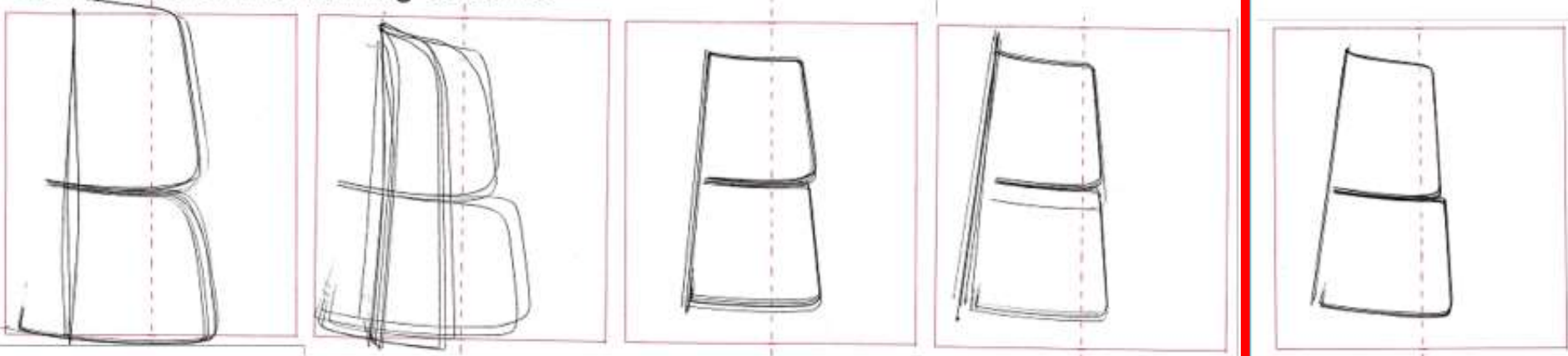
境野従来手法

提案自己回帰なし

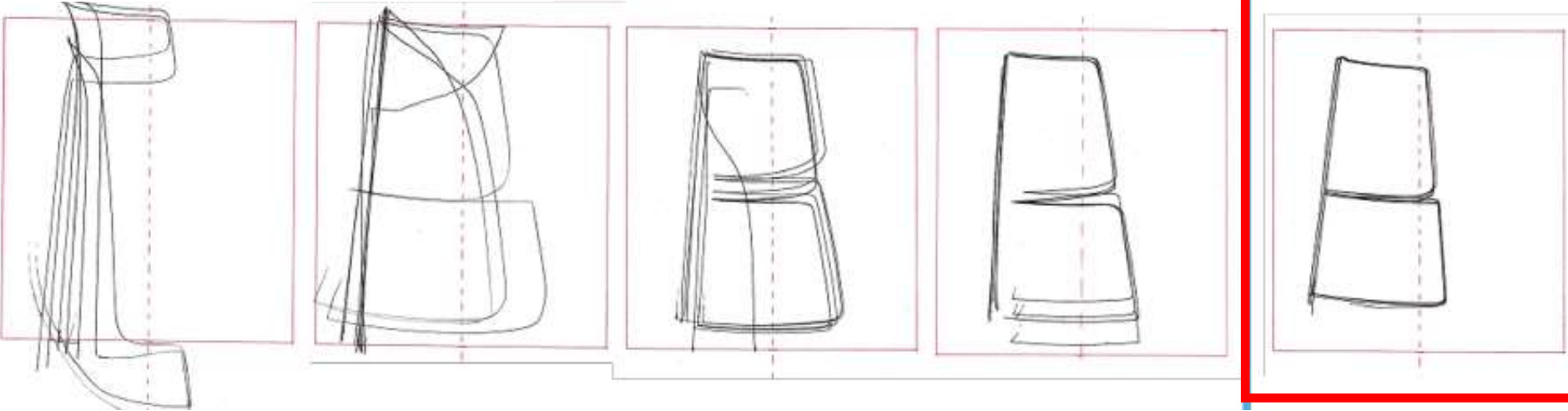
提案自己回帰あり

字を書くタスクによる実証

教師データ: 40文字



教師データ: 4文字



従来手法1

従来手法2

境野従来手法

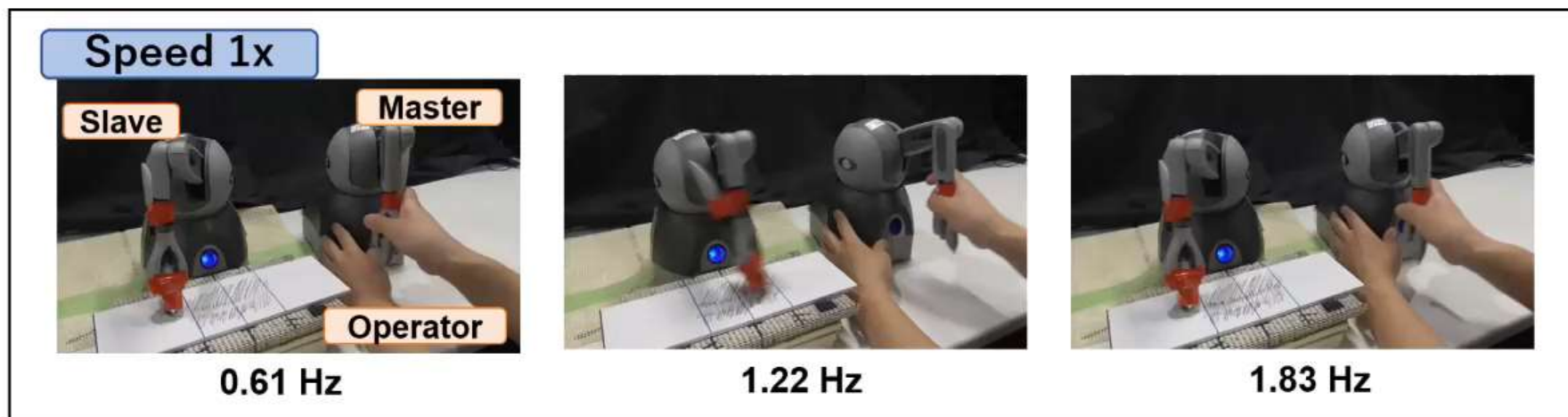
提案自己回帰なし

提案自己回帰あり

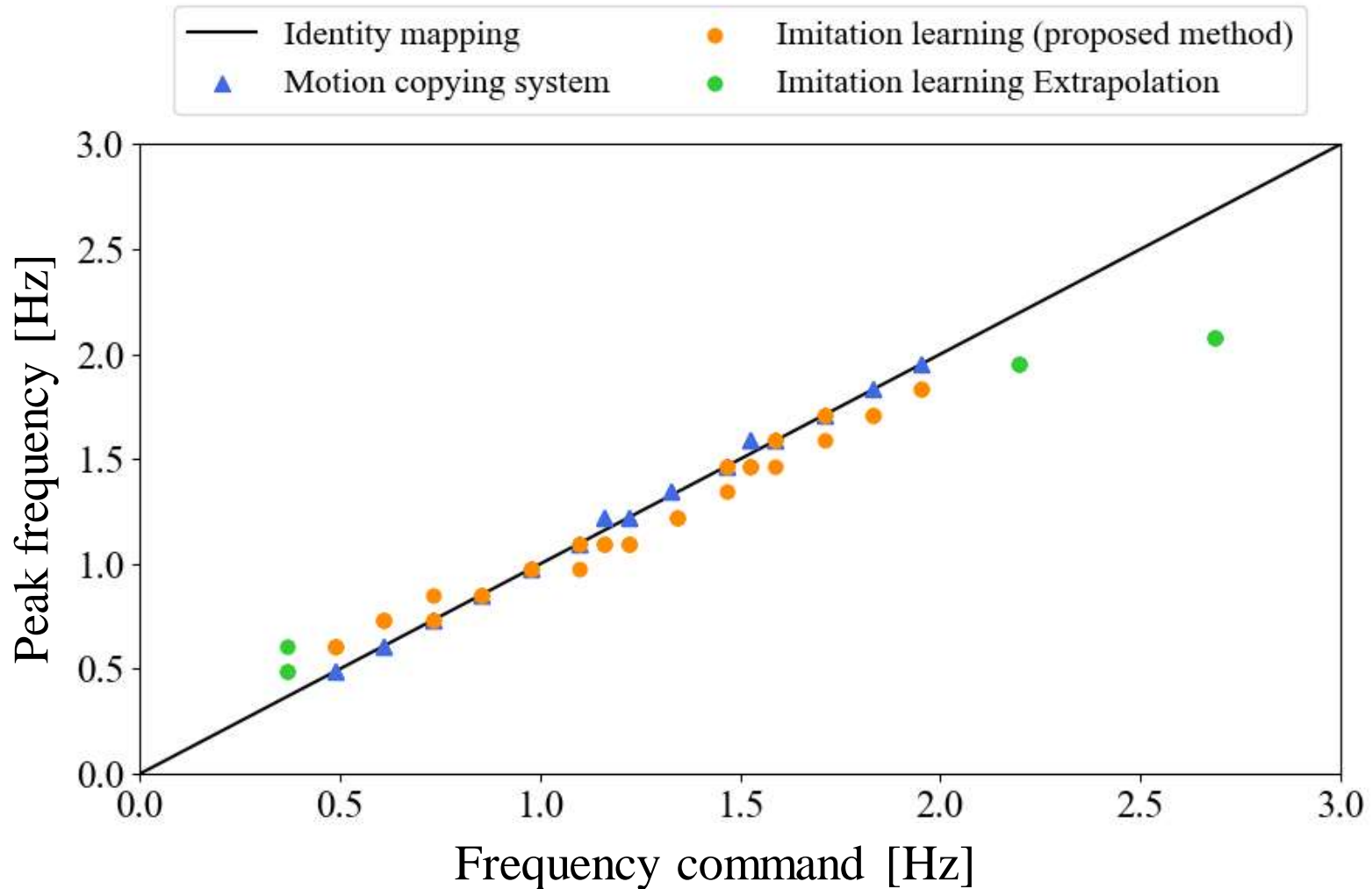
可変速動作生成 (時間方向の汎化性能の獲得)

可変速な動作生成手法

- ① 複数の速度での動作データを学習

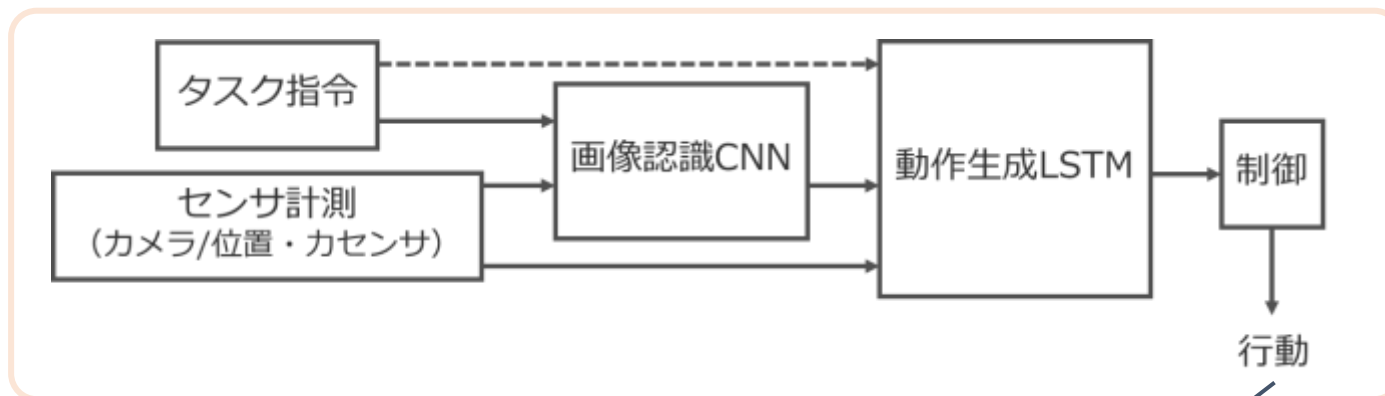


指令周波数と実際の応答周波数



• 直列型モデル

各モジュールが直列で動作



人間並みの速い動作を実現したい
速いサンプリングを実現する必要がある

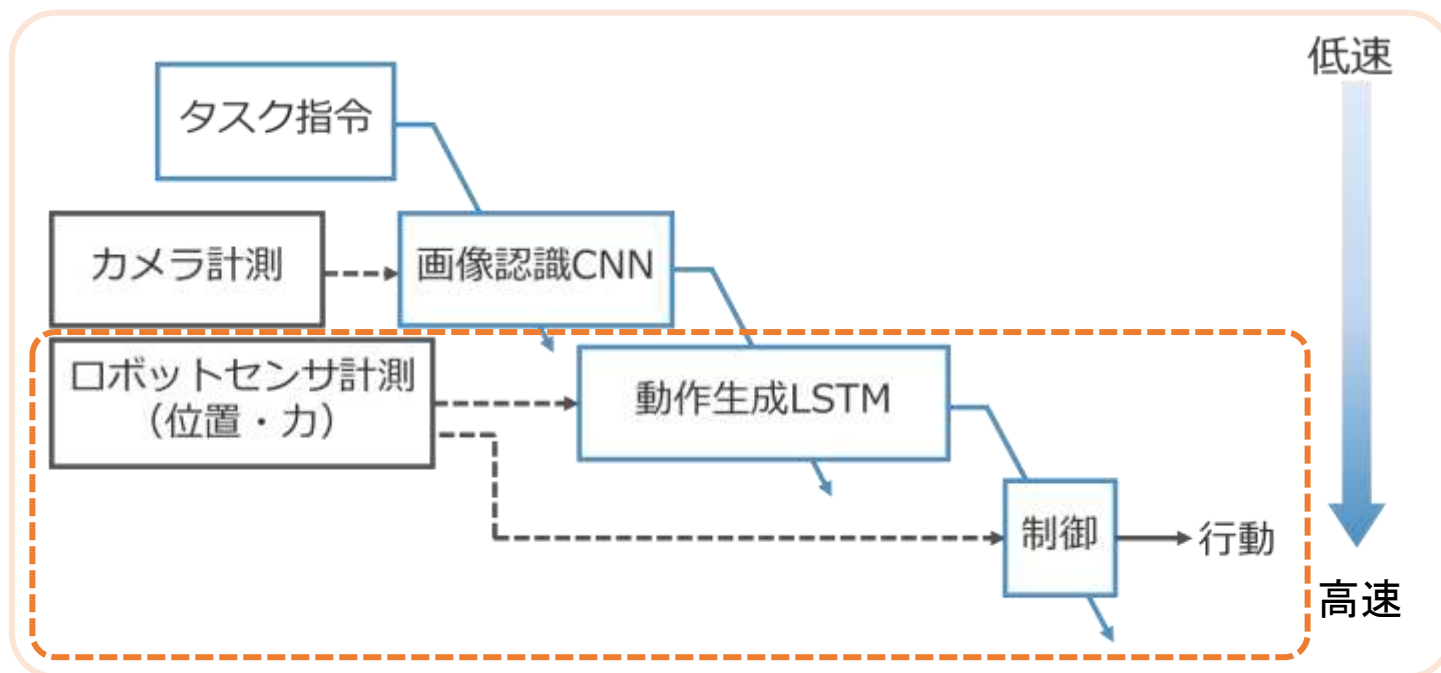
実現できる仮定:

- 全てのモジュールが高速に動作する
- マシンパワーが最強(センサ計測の速度・NNの推論速度)
- 学習データが無限にあり十分に学習できる

階層化動作生成

- 階層型モデル

各モジュールが時間的に並列で動作

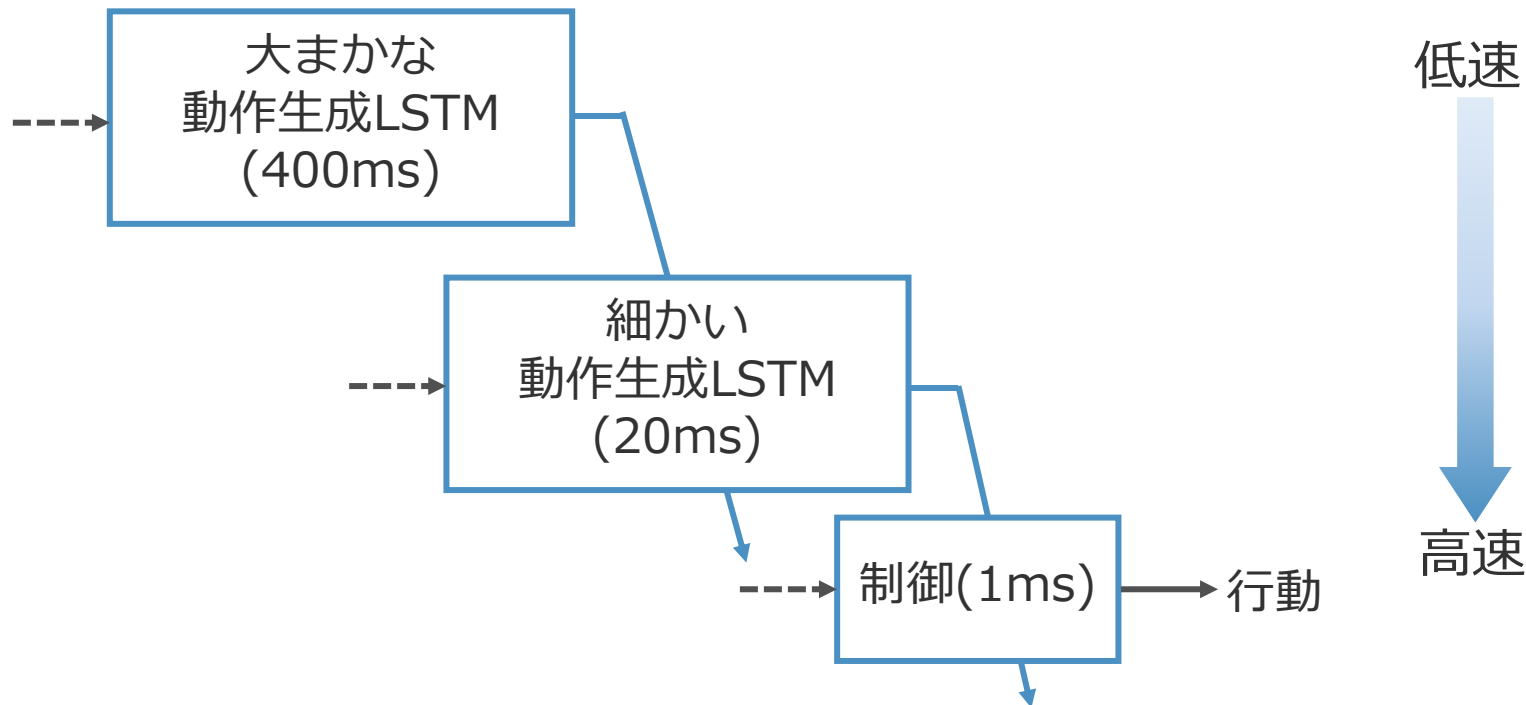


階層化動作生成

階層型の動作生成

目標：

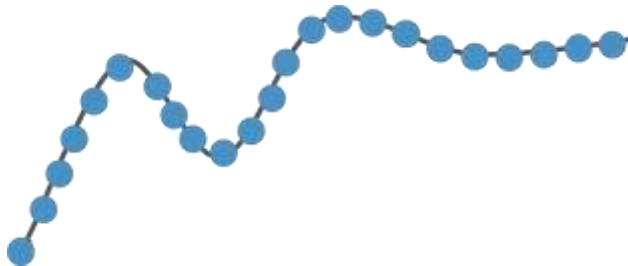
1. 長周期タスク（複雑な行動サイクル）への対処
2. 学習の効率化
3. （タスク間の汎化）



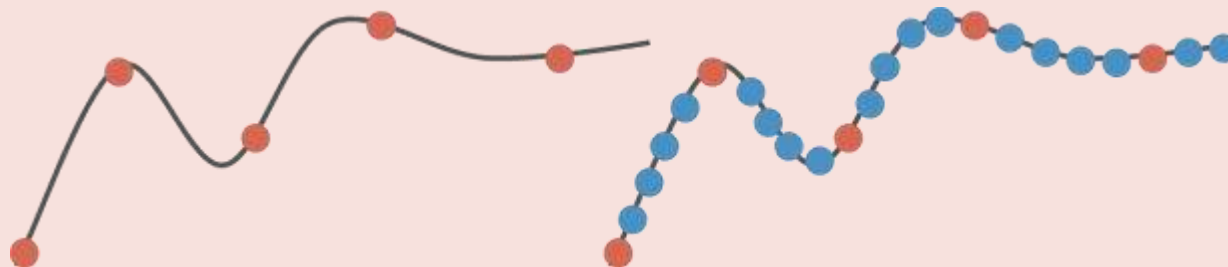
階層化動作生成

階層型の動作生成

- 従来モデル



- 提案モデル（階層型）



上位層：サブゴールを設定

下位層：サブゴールに対して動作

長周期タスクの実現（上位・下位層でタスクが明確化）

- それぞれのLSTMが学習すべき問題が簡単になる
- 予測された未来の状態に追従するように下位層は動作生成可能

階層化動作生成

階層型の動作生成

検証タスク：“ABC”を書く

タスク1回完了時間：15s（LSTMの推論：750step）



学習データ

- 高さ：3段階
(25mm,45mm,70mm)
- 1データ（約1分）あたり
約4回“ABC”を連続して書く

階層化動作生成

深層型と階層型の比較

結果： 深層型・階層型ともにすべて成功

学習時間

深層型： 11層（40時間～52時間）
5層（失敗）

階層型： 上位層-5層（42分）
下位-5層（3時間）

さらに短い時間でも成功

まとめ

- 力をも模倣学習することで、
環境の変動や人間との共同が可能なAIを獲得
- 制御の遅れを考慮した動作収集・模倣学習ができるため
人間並みの高速動作が可能
- 長期間の動作生成を安定化可能
- 教示動作よりも早く動作させることが可能
- 認識・判断、動作生成、制御を独立に設計可能

- タスク専用のハードウェアが不要ということはそもそもロボットを買う必要がなくなる。
- 必要なときに必要なだけロボットを借りればよい。

Robot as a service

ロボットシェアリングの時代へ

3軸ロボットにおける単純なタスクは自律化済み

課題

- 画像情報との統合
高速画像処理や物体認識等が必要になる可能性が高い
- 力制御を実現可能な多自由度汎用ロボットの製作
ハンドロボットは製作中
アクチュエーション技術も所有

不定形物の操作や力加減が必要な技能の自律化に有効

- ex) ピッキング、バリ取り、研磨作業、左官、ペンキ塗り、組立作業、果物の採取、箱詰め作業、料理、洗濯、アイロンがけ、掃除
- 画像処理技術やロボット製作技術を持つ企業との共同研究を希望。

本技術に関する知的財産権



- 発明の名称：情報処理装置、方法、及びプログラム
- 出願番号：特願2020-120740
- 出願人：筑波大学
- 発明者：境野翔

お問い合わせ先

- 筑波大学
- システム情報エリア支援室（研究支援）
- Tel: 029-853-4989
- E-mail: sien3@un.tsukuba.ac.jp