

ロボット制御における Attention 機構を用いたモデルの実験的評価

Experimental Evaluation of an Attention-Based Model for Robot Control

○田中大介*, 西原優空*

○ Daisuke Tanaka*, Yura Nishihara*

*新居浜工業高等専門学校

*National Institute of Technology (KOSEN), Niihama College

キーワード： 非線形同定 (nonlinear system identification),
産業用ロボットアーム (industrial robot arm), Cross-Attention, Perceiver IO,
ニューラルネットワーク (neural networks)

連絡先： 〒 792-8580 愛媛県新居浜市八雲町 7-1
田中大介, E-mail: d.tanaka@niihama-nct.ac.jp

1. はじめに

産業用ロボットの性能向上には、高精度な予測モデルが不可欠であり、軌道シミュレーション、Model Predictive Control (MPC)、フィードフォワード制御などにおいて中心的な役割を担う。近年、機械学習とりわけ Attention 機構を基盤とした手法が、動的システムに内在する複雑な時間的・空間的関係の抽出に有望であることが示されている¹⁾。本研究では、高次元かつ不均質なデータを扱うための Cross Attention 機構を有する Perceiver IO²⁾ を取り上げ、その有効性を検証する。Self Attention の有効性は広く検証されている一方で、Cross Attention がもたらす固有の利点については十分に解明されていない。そこで本稿では、Forward モデル（トルクから関節角の予測）と Inverse モデル（目標関節状態から必要トルクの推定）の同定問題

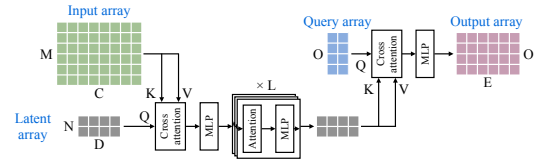


Fig. 1 Overview of Perceiver IO

における Cross Attention の有効性について検証した結果を報告する。

2. Perceiver IO の概要

Perceiver IO は、Transformer 系の機械学習モデル（回帰モデル）であり、高次元かつ不均質なデータを効率的に扱う柔軟なアーキテクチャである。構造は図 1 に示す通り、概ね以下の三段から成る。(i) 入力配列の特徴を Cross Attention により学習可能な潜在配列（latent array）に写

像する．(ii) 潜在配列を Self Attention と多層パーセプトロン (MLP) により更新し，ベクトル間の関係を捉えるとともに，時系列データの時間的特性を間接的に保持する．(iii) 潜在配列を，所望長の出力クエリ配列 (query array) へ Cross Attention で写像し，入力から必要情報を抽出して任意形式の出力を得る．この構造により，入力全体の情報を効率よく圧縮しつつ，時系列に内在する時間関係を考慮でき，多入出力かつ複雑な動的システムに対して有効である．

3. 実験

3.1 データセットとタスク

本研究では，KUKA KR300 R2500 ultra SE ロボットから収集された Industrial Robot ベンチマークデータセット³⁾を用いる．本データセットは 10 Hz でサンプリングされた 43,622 ステップからなり，各ステップには 6 軸のモータトルク $\tau \in \mathbb{R}^6$ (単位: Nm) と 6 軸の関節角 $q \in \mathbb{R}^6$ (単位: deg.) それぞれの数値データが含まれる．

Forward モデルはモータトルクから関節状態 (例: 関節角) を予測する．本研究では，次式の 1 ステップ予測を仮定する．

$$q(t) = f(q(t-1), q(t-2), \tau(t-1), \tau(t-2)) \quad (1)$$

Inverse モデルは制御タスクにおいて重要であり，所望軌道や目標状態を実現するために必要なモータトルクを与える．

$$\tau(t) = g(q(t), \dot{q}(t), \ddot{q}(t)) \quad (2)$$

3.2 モデル構成と評価指標

本研究では，Forward モデルおよび Inverse モデルの同定に対して，全結合ニューラルネットワーク (NN) と Perceiver IO の 2 種類を比較対象とした．NN は単純なベースラインとして構築し，入力層・出力層の間に 1 層の隠れ層 (ユ

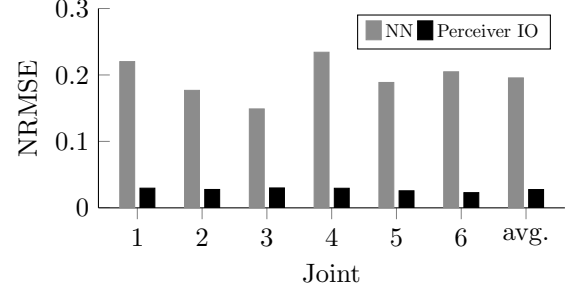


Fig. 2 NRMSE of forward model for test data

ニット数 100) を持つ構成とした．Forward モデルにおいては，式 (1) で示した 24 次元の入力を 6 次元出力に写像する構成とし，Inverse モデルでは式 (2) に基づく 18 次元の入力を 6 次元出力に写像する構成とした．

Perceiver IO については，Forward モデルでは 2 時刻分のトルクおよび関節角を連結し，12 次元の入力系列としてモデルに与え，6 次元の関節角を出力するように設定した．Inverse モデルでは単一時刻の 18 次元入力から 6 次元のトルクを出力するように設定した．データセットのうち，学習データの 80% を訓練に，残り 20% を検証に用い，両モデルのハイパーパラメータは同程度の計算コストとなるように調整した．

なお，モデルの性能評価には正規化二乗平均平方根誤差 (NRMSE) を用いることとした．

3.3 結果

図 2 および図 3 に Forward モデルと Inverse モデルそれぞれにおける比較結果を示す．Forward モデルでは全軸に対して Perceiver IO が NN を上回る精度を示した．また，Inverse Model においては一部 NN の方が精度が高い軸があったものの，特に第 6 軸においては Perceiver IO が高い性能を発揮した．

これらの結果から，Perceiver IO は多入出力構造および非線形動特性を持つシステムの同定に対して有効であり，従来の全結合 NN と比較

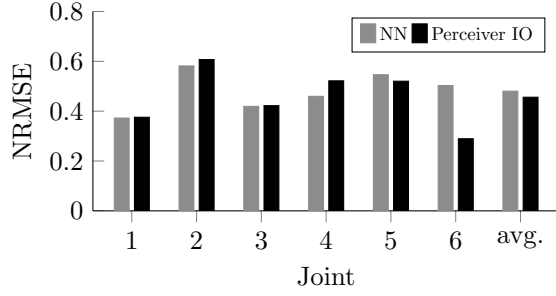


Fig. 3 NRMSE of inverse model for test data

して汎化性能とロバスト性の両立が可能であることが確認された。これにより Perceiver IO の制御モデルへの応用可能性が示された。

4. 結論

本研究では、産業用ロボットの Forward モデル・Inverse モデルの同定問題に対し、Perceiver IO と NN を比較した。結果として、Perceiver IO は両タスクで高い性能を示し、複雑な非線形ダイナミクスの扱いに適していることが示唆された。今後の課題として、他の動的システムの同定手法との比較や、実時間制御への適用、他の機械学習手法との統合の検討が挙げられる。

謝辞

本研究の一部は JSPS 科研費 JP23K28369 ならびに JP22K17918 の助成を受けた。

参考文献

- 1) Javier Fañanás-Anaya, Gonzalo López-Nicolás and Carlos Sagiús: Dynamical system simulation with attention and recurrent neural networks, Neural Computing and Applications, 2024.
- 2) Andrew Jaegle, Sebastian Borgeaud, Jean-Baptiste Alayrac, Carl Doersch, Catalin Ionescu, David Ding, Skanda Koppula, Daniel Zoran, Andrew Brock, Evan Shelhamer, Olivier Hénaff, Matthew M. Botvinick, Andrew Zisserman, Oriol Vinyals and João Carreira: Perceiver IO: A General Architecture

for Structured Inputs & Outputs, International Conference on Learning Representations, 2022.

- 3) Jonas Weigand, Julian Götz, Jonas Ulmen and Martin Ruskowski: Dataset and Baseline for an Industrial Robot Identification Benchmark, 6th Edition of the Workshop on Nonlinear System Identification Benchmarks, Leuven, Belgium, 2022.