T2R2 東京科学大学 リサーチリポジトリ
Science Tokyo Research Repository

論文 / 著書情報
Article / Book Information

題目(和文)	マルチモーダル情報を用いたロボットコントロールシステムのための シンプルでロバストなアーキテクチャの提案と実ロボットにおける動作検証
Title(English)	
著者(和文)	秋川元宏
Author(English)	Motohiro Akikawa
出典(和文)	学位:博士(工学), 学位授与機関:東京工業大学, 報告番号:甲第11735号, 授与年月日:2022年3月26日, 学位の種別:課程博士, 審査員:山村 雅幸,小野 功,青西 亨,瀧ノ上 正浩,長谷川 晶一
Citation(English)	Degree:Doctor (Engineering), Conferring organization: Tokyo Institute of Technology, Report number:甲第11735号, Conferred date:2022/3/26, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Category(English)	Doctoral Thesis
種別(和文)	論文要旨
Type(English)	Summary

Powered by T2R2 (Science Tokyo Research Repository)

（論文博士）
(Dissertation Doctorate)

論　文　要　旨　（和文2000字程度）

Dissertation Summary (approx. 2000 characters in Japanese)

報告番号 For administrative use only	乙 第　　号	氏　名 Name	秋川　元宏

（ 要　旨 ）
(Summary)

人間と相互作用するロボットが開発されている。このようなロボットの制御システムは長い年月をかけて研究されている。古典的な制御方法では、システムの設計者がプログラムとして作り込むことでロボットの制御をしていた。近年では深層学習を用いた制御システムの開発が主流である。深層学習を用いた制御システムの中で状況認識の高精度化を目論み、マルチモーダル情報を入力とするものもみられる。しかし、マルチモーダル情報を入力とする深層学習を用いた制御システムには2つの欠点がある。第一に計算コスト、第二にノイズ耐性である。マルチモーダル情報を扱う深層学習はネットワークの規模が大きくなることで、計算コストが非常に高くなる。ネットワークの一部の結線を除去することで計算コストの増加を抑える研究も存在するが、どの結線を除去するかは自明ではなく、システム設計者に対し大きな負担を強いることになっている。ロボットは実環境において使役されるが、光源の位置が少し移動しただけで不安定になるになるなどの実環境由来の問題点が指摘されている。光源の位置のずれなどは環境ノイズと考えることができる。本研究では上記の2点を克服するロボット制御用アーキテクチャを提案する。アーキテクチャの提案に際し、アーキテクチャの特性を明らかにするために、擬似乱数によって生成された疑似データを用い学習および動作検証を行なった。さらに、リーチングタスクを行うアームロボットから得た実データを用いて学習実験を行い、実データに対しても提案アーキテクチャが動作可能であることを示した。

提案アーキテクチャは3つのモジュールで構成されている。最も重要なモジュールはアソシエイタユニットで、単一のホップフィールドネットワークで構成されている。このアソシエイタでマルチモーダル情報を統合し、記憶／想起を行う。もう１つのモジュールはエンコーダユニットである。エンコーダユニットは情報ごとに独立したエンコーダを持ち、3層フィードフォワードニューラルネットワークで構成されている。本研究では画像、音声、アクチュエータの位置の３種類のマルチモーダル情報を統合するため、３つの独立したエンコーダがエンコーダユニットに存在することとなる。これらのエンコーダはセンサーから得たアナログ情報をホップフィールドネットワークが処理できるバイナリー情報へ変換する符号器として機能することとなる。最後のユニットがデコーダユニットである。デコーダユニットもエンコーダユニット同様に情報の種類ごとに独立した3層のフィードフォワードニューラルネットワークで構成されている。デコーダユニットはアソシエイタで処理されたバイナリー情報をシステムの出力となるようにアナログ情報への変換を行う復号器として機能する。学習において、エンコーダとデコーダは情報の種類ごとのオートエンコーダとして同時に学習される。

擬似データを用いたアーキテクチャの動作検証実験において、提案アーキテクチャは単体でマルチ

モーダル情報を処理するオートエンコーダに比べ遥かに高いノイズ耐性を示した。また、提案アーキテクチャの高いノイズ耐性能力はアーキテクチャのサイズにより変化し、アーキテクチャのサイズが大きくなるにつれノイズ耐性が高くなることが判明した。

リーチングタスクを行うアームロボットから得た実データを用いた学習実験を行うために、動作検証実験で用いたアーキテクチャに変更を加えた。最も重要な変更点は、ホップフィールドネットワークのサイズ大きくし、時刻tの状態および時刻t+1の状態を同時に銘記と想起できるようにしたことである。これにより、動作検証実験のアーキテクチャでは未実装であった、制御信号の生成というロボット制御システムにおける重要な機能を追加することができる。

実データを用いた学習実験では、3方向、中間9状態のリーチングタスクを学習し、すべての方向へリーチングできることを示した。中間状態においても、軌跡およびシステムからの出力をアームロボットで再生したときの動作を目視で観察し、9状態のうち、8状態において、理想出力と遜色なく動作していることを確認した。

提案アーキテクチャは、オートエンコーダのように単一アーキテクチャでマルチモーダル情報を処理するアプローチと比べて、独立したエンコーダ、デコーダと計算コストの小さいホップフィールドネットワークで構成されていることから計算コストの削減に大きく貢献している。ノイズ除去能力の高いホップフィールドネットワークをアソシエイタとして採用することで、高いノイズ耐性を得ることができた。ただし、ホップフィールドネットワーク単体で扱った時に比べ、ノイズ除去能力は落ちている。ホップフィールドネットワークに銘記するパターンはエンコーダで自律的に生成されるため、必ずしも理想的なパターンとなっていないからだと考えられる。このように提案アーキテクチャはマルチモーダル情報を用いるロボット制御システムの基礎としては十分な能力を有するが、ロボットのリアルタイム制御を行うためには、エンコーダの改良が必要と考えられる。

備考：論文要旨は、和文2000字と英文300語を1部ずつ提出するか、もしくは英文800語を1部提出してください。
Note: Dissertation summaries must be written in either of the following formats: (A) both in Japanese (approx. 2000 characters) and in English (approx. 300 words), or (B) in English (approx. 800 words).
注意：論文要旨は、東工大リサーチリポジトリ（T2R2）にてインターネット公表されますので、公表可能な範囲の内容で作成してください。
Important: Dissertation summaries will be published online on the Tokyo Tech Research Repository (T2R2). Do not include information treated as confidential under certain circumstances.