

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),<br>学位授与機関:東京工業大学,<br>報告番号:甲第11735号,<br>授与年月日:2022年3月26日,<br>学位の種別:課程博士,<br>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),<br>Conferring organization: Tokyo Institute of Technology,<br>Report number:甲第11735号,<br>Conferred date:2022/3/26,<br>Degree Type:Course doctor,<br>Examiner:,,,,, |
| 学位種別(和文)          | 博士論文  |
| Category(English) | Doctoral Thesis   |
| 種別(和文)            | 審査の要旨   |
| Type(English)     | Exam Summary  |

## 論文審査の要旨及び審査員

(2000字程度)

| 報告番号  | 乙 第 号    | 学位申請者 | 秋川 元宏  |     |
|-------|----------|-------|--------|-----|
|       | 氏 名      | 職 名   | 氏 名    | 職 名 |
| 論文審査員 | 主査 山村 雅幸 | 教授    | 長谷川 晶一 | 准教授 |
|       | 小野 功     | 准教授   |        |     |
|       | 青西 亨     | 准教授   |        |     |
|       | 瀧ノ上 正浩   | 准教授   |        |     |

本論文は「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」と題し全5章からなる。

第1章「ロボット制御における技術」では背景として人間と相互作用するロボットが開発されていること、およびロボット制御方法の変遷について述べている。古典的な制御方法では、システム的设计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と述べている。近年ではマルチモーダル情報を入力とした深層学習を用いた制御システムの開発が主流であると述べている。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つの欠点があることを指摘している。第一に計算コスト、第二にノイズ耐性であるとしている。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としている。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と指摘している。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などの実環境由来の問題点が既存研究にて指摘されていると述べている。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と述べている。上記の2つの欠点を克服するロボット制御用アーキテクチャの提案が本研究の目的であると述べている。

第2章「マルチモーダル情報を統合するアソシエイティブメモリを含むシステムの提案」では、提案アーキテクチャの概要と最適化手法について次のように述べている。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種類の情報を統合することと仮定し、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情報へ変換する符号器として機能する。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。

第3章「提案手法における性能と挙動検証」では、提案アーキテクチャに対し、擬似データを用いた動作検証実験を行なっている。結果として、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高いノイズ耐性を示したとしている。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としている。また、提案アーキテクチャにおいて、第1章で述べた問題であるコストとノイズ耐性の観点で議論を行なっている。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と主張している。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としている。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と述べている。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としている。

第4章「提案手法の改良と実データにおける動作検証」では、提案アーキテクチャの改良を行うとともに、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た学習実験を行なったと述べている。最も重要な改良点は、ホップフィールドネットワークのサイズを大きくし、時刻 $t$ の状態および時刻 $t+1$ の状態を同時に銘記/想起できるようにしたことであるとしている。これにより、第2章で提案された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としている。実データを用いた学習実験では、3方向、中間状態を含め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と主張している。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としている。

第5章「総合討論」では、本研究の結果をまとめ、総合的な討論を行なっている。従来の深層学習では学習点を増やし精度の向上を図っており、学習点以外での出力は結果が保証されないことを指摘している。提案アーキテクチャは深層学習の手法を使いながら、記憶ベースの構造を含むことで、高い認識力と安定した挙動を示すことを議論している。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であることを指摘している。この指摘に対し、エンコーダとなり得る候補について、学習における前処理について議論している。

以上を要するに、マルチモーダル情報を用いる次世代の知能ロボットのために、記憶のアーキテクチャを提案してその基本性能を明らかにした上で、実機においてその効果をデモンストレーションしたもので、工学上貢献するところが大きい。よって、博士（工学）の学位論文としてふさわしいものと認められる。

注意：「論文審査の要旨及び審査員」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。