

知的情報処理特論

2015/04/16

担当：伊藤

T15M074

「ディープニューラルネットワークを用いたロボットの行動のマルチモーダル統合学習」

『Multimodal integration learning of robot behavior using deep neural networks』

-Kuniaki Noda , Hiroaki Arie ,Yuki Suga ,Tetsu Ogata

●内容

- ① 本文概略
 - ・ HIGHLIGHTS
 - ・ ABSTRACT
 - ・ 1.Introduce
- ② 用語

①概略

それぞれの概略を付随する。Introduction では段落ごとに区切っている。

・ HIGHLIGHTS

- 感覚運動統合学習のための新規の計算フレームワーク。
- ディープオートエンコーダを利用し、クロスモーダルメモリ検索。
- 獲得したマルチモーダルな機能を利用し、ノイズに頑健な行動認識。
- マルチモーダル因果関係の取得と感覚運動予測。

・ ABSTRACT

人間が正確に周囲の世界を理解するにはマルチモーダル統合が必要不可欠である。これを再現できればロボットの実用化に貢献することができるが、従来の機械学習アルゴリズムには問題があり、知覚運動情報処理プロセスはモーダル依存によって達成されていた。本稿では生の RGB 画像、音声スペクトル、及び関節角度を含む感覚運動時系列データの統合とディープラーニングからアプローチすることでマルチモーダル融合した新しい計算の枠組みを提案する。次の実験と実証も行った。(1)ディープオートエンコーダの情報補完機能を利用し、クロスモーダルメモリ検索。(2)マルチモーダルな機能の汎化能力を利用し、ノイズに強固な行動認識を行う。獲得した因果関係に基づいており、(3)マルチモーダル因果関係の

取得と感覚運動予測を行う。

・ 1.Introduce

人間は複数のソースから外部情報をとらえるが、これにはマルチモーダル統合が貢献しており、認知科学の研究はその仕組みを明らかにした。さらにクロスモーダルグルーピングは知覚に重要な役割を果たしており、人間のマルチモーダル統合学習を再現することは洗練されたロボットの知能を実現させる。

問題が残る関連研究[7,8,9]が報告されているがマルチモーダル統合はロボット工学において困難な問題であった[5,6]。いくつかの試みはあったが、ほとんどの因果モデルは特定の部分に注目しており、様式が限られている。[10,11,12]

【以下引用】

[5] 知性の代替エッセンス

[6] マルチモーダル統合の生物学的見解

[7] マルチモーダル統合と空間的な開発の単一表現のための異なるメカニズム

[8] 実際のロボットのナビゲーション作業用のグリッドセルとビジュアルプレイスセルのマルチモーダル統合

[9] 生物学的にマルチモーダル統合に影響を与える：人間とロボットの相互作用ゲームの干渉において

[10] AI ロボットの紹介

[11] シンプルな視覚運動マッピングと模倣に基づいた人間のパートナーとの相互作用ルール学習

[12] 感覚運動因果関係スペースの圧縮を通じて相互作用のルールを学ぶ

知覚精度の高い最先端のデバイスを用いて、利用可能な感覚情報の量での改善は顕著だったが、単一の表現にマルチモーダル感覚入力を融合することによって、強固な挙動制御と環境認識が実現するような拡張性のある画期的な計算モデルは、まだ提案されていない。

大規模データの処理に対する計算モデルについては、ディープラーニングアプローチが近年大きな注目を集めており大きな業績も残しているが、静的なものが主である。そのため本論文では動的なものについても検討した。

今回の研究では認知科学の研究によって報告された基本原理を、計算モデルの構築により実用化することであり、設定した機能の実証について実験を行います。

提案されたフレームワークはディープオートエンコーダと時間遅延ニューラルネットワークの変形を組み合わせ、よりレベルの高いマルチモーダルな特徴を自己組織化する。さらに内部部分的に入力データを外部からマスクする。オートエンコーダの特徴を利用することで、入力と出力のモデルの ID マッピングを行う。

二つの実験を行うことで、提案したフレームワークを評価する。一つ目の実験では、ダイレクトティーチングによって生成された人型ロボットの 6 つの異なるオブジェクト操作行動を用いて提案モデルを訓練し、関節角度情報を利用して行動認識性能の強固性および信頼性を顕著に向上させることを実証した。

二つ目の実験では画像・関節角度・音声信号を取り込むことでマルチモーダル統合学習を拡張した。ベルリンギングタスクを設計し、3つのモダリティで構成された知覚運動シーケンスを利用する提案モデルを訓練した。結果、メモリ検索に成功し、因果依存関係を理解した。

このセクションに加えて、以下のセクションで論文は構成されています。

2 節では 私たちは簡単に深いネットワークを訓練するためのヘッセのない最適化を検討。

3 節では 私たちはマルチモーダル時系列学習の一般的な枠組みを説明します。

4 節では 私たちは人型ロボットを利用しオブジェクト操作行動のマルチモーダル統合学習によって私たちの提案モデルの実用性と有効性を提示する。

次に、5 節に 私たちの提案モデルはその方法を実証した。ベルリンギングタスクに移動します。

6 節では我々は前作との関係で、検討し私たちの提案フレームワークと結果を議論する。

最後に、我々は、7 セクションで私たちの仕事を終える

●用語

・ディープオートエンコーダ

中間層 1 層だけ作り、出力層を取り除き、中間層を入力層と見なしてもう 1 層積み上げる。以上を繰り返してニューラルネットワーク作る方法

・クロスモーダル

複数の感覚を統合することによって、単一の感覚刺激では困難であった感覚や体験を呈示する方法

- ・ディープラーニング

機械学習の一種で、ニューラルネットを何層も重ねたものを用いてクラス分類や回帰を行うための手法。複数の入力値を元に全体や細部を見て全体の認識をしないおすアプローチを採用している。

- ・教育なし学習

機械学習の手法の一つで、出力すべきものがあらかじめ決まっていないという点が教師あり学習と大きく異なる。

- ・