

令和元年6月22日現在

機関番号：25301

研究種目：基盤研究(C) (一般)

研究期間：2015～2018

課題番号：15K00244

研究課題名(和文) ロボットによるマルチモーダル言語獲得 格助詞の学習

研究課題名(英文) Multimodal Language Acquisition by Robots - Learning Case Particles

研究代表者

岩橋 直人 (Iwahashi, Naoto)

岡山県立大学・情報工学部・教授

研究者番号：90394999

交付決定額(研究期間全体)：(直接経費) 3,600,000円

研究成果の概要(和文)：本研究では、従来可能であった内容語と動詞のみからなる単純な物体操作指示発話を
用いた学習を、機能語である格助詞を含んだ文にまで拡張することが目的であった。

動作と発話(格助詞を含む)からなる人間の行為に対して、ロボットが、動作と発話(格助詞を含む)で応答
することを学習するニューラルネットワークモデルの開発に成功した。本手法におけるオリジナルな特徴は次の
三つである：(1)人間の動作、発話、状況画像の情報を入力。(2)ロボットの動作、発話の情報を出力。
(3)複数種類の発話行為に対して適応的に応答可能。

さらに、このモデルを用いて、統合失調症患者の言語コミュニケーションのモデル化にも成功した。

研究成果の学術的意義や社会的意義

ロボットが、人間とのマルチモーダルなコミュニケーションを通して、実世界の事物に直接グラウンドしない機
能語である格助詞とその文法上の機能を学習する手法は極めて独創的である。

動作と発話からなる人間の行為に対して、ロボットが、動作と発話で応答できるニューラルネットワークモデル
を世界に先駆けて開発したことは、日本のディープラーニングのロボティクスおよび言語処理応用研究における
高い競争力の維持に大いに貢献している。

さらに、精神科医療に対して、人工知能研究が貢献する道筋をつけることができたことは医学的にも大きな意
義がある。

研究成果の概要(英文)：The purpose of this research was to extend the learning using simple object
manipulation instruction utterance consisting only of content words and verbs, which was
conventionally possible, to sentences including case particles that are functional words.

We have successfully developed a neural network model that learns that robots respond to motion and
speech (including case particles) with motion and speech (including case particles). The original
features in this method are the following three: (1) Input information of human motion, speech, and
situation image. (2) Output information on robot operation and speech. (3) Adaptively respond to
multiple types of speech acts.

Furthermore, using this model, we also succeeded in modeling verbal communication in schizophrenic
patients.

研究分野：人工知能

キーワード：記号創発ロボティクス ヒューマンロボットインタラクション 機械学習

1. 研究開始当初の背景

日常支援ロボットでは、ユーザとのコミュニケーション機能が極めて重要である。特に、ユーザが音声で行動の指示を与えることができる対話機能は、複雑な指示操作を苦手とする多くのユーザにとって必要不可欠な機能である。しかしながら、現状のロボット対話処理技術は必要なレベルにまったく到達していない。このことが、日常生活支援ロボットの普及の大きな障壁となっている。

ユーザが生活支援ロボットと安心してコミュニケーションをするためには、ロボットが、ユーザの生活空間や生活習慣に応じて、ユーザの曖昧な指示発話の意味を適切に理解して行動できる、マルチモーダルな語用論的言語理解の機能が必要である。ところが、従来の言語処理技術では、あらかじめ固定的で閉じた記号系が与えられ、記号の持つ意味は感覚運動情報とは無関係にその内部で循環的に定義されているため、この機能は実現できない。

この問題を解決するために、申請者は、対話と行動のコミュニケーション能力を、予めロボットに与えるのではなく、ロボット自身が、実世界における音声と視覚と動作による人間とのコミュニケーションを通して学習する言語獲得計算機構 L-Core を既に開発済である。L-Core は、言語的知識、行動に関する非言語的知識、実世界知識などの種々の知識を適応的に関連付け、状況に応じてユーザの意図を適切に推定し、発話や動作を適切に理解・生成することが可能であり、そのようなマルチモーダルな語用論的言語能力の獲得が計算論的に構成可能であることを示していた。

2. 研究の目的

上記の言語獲得計算機構 L-Core は、実世界に直接グラウンドする内容語である形容詞と名詞と動詞のみからなる単純な物体操作指示発話（例、「カーミット赤い箱乗せて」）を学習の対象としていた。本研究では、人間の言語獲得の構成論的な理解を深めるために、L-Core の学習対象を、実世界に直接はグラウンドしない機能語である格助詞を含んだ文にまで拡張することが目的であった。具体的には、「カーミットを赤い箱に乗せて」や「小さいボールを大きいボールから遠ざけて」といった格助詞を含むより自然な物体操作指示発話を学習対象として、格助詞の知識を含んだマルチモーダル言語モデルを学習する機能の実現を目指すものであった。

3. 研究の方法

研究期間4年間のうち、最初の二年間で機能語を含む物体操作指示発話と物体操作情報を相互に対応付ける機械学習手法を開発した。引き続き三年目で、実際のヒューマンロボットコミュニケーションで利用可能にものとして、この手法を基盤にして、人間の動作と発話からなる行為に対して、ロボットが動作と発話からなる行為を生成するための機械学習手法を開発した。さらに4年目で、人間の語用論的言語認知機能の構成論的な理解を深めるために、この機械学習手法を基盤にして、精神障害の計算論的モデルを構築した。

4. 研究成果

計画の最初の二年間（27、28年度）で、機能語を含む物体操作指示発話と物体操作情報の対応付けを、ディープラーニングにより End-to-end で学習する手法の開発に成功した。提案手法は、深層格構造の解析における物体クラス認識の計算過程と、表層格構造を解析する計算過程、および表層格構造と深層格構造の対応付ける計算過程を陽に設けることなく、全体の計算過程の中に統合して埋め込み、全体性能を最適化するものである。

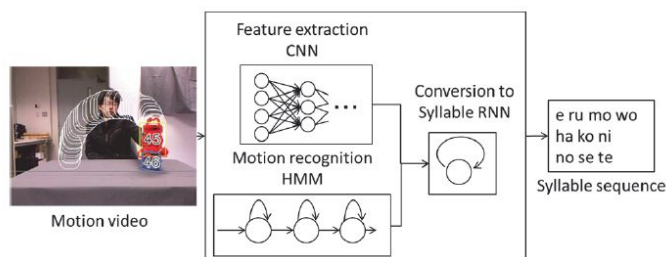


図 1 物体操作指示発話と物体操作情報を対応づけるニューラルネットワークモデル

これに加えて、引き続き三年目（29年度）は、動作と発話（格助詞を含む）からなる人間の行為に対して、ロボットが適切に、動作と発話（格助詞を含む）からなる行為で応答することを学習する、ヒューマン・ロボット・マルチモーダル・コミュニケーションのニューラルネットワークモデル（MCNNM: Multimodal Communication Neural Network Model）を開発した。従来のいかなる手法でも実現されていない、本手法におけるオリジナルな特徴は次の三つである：（1）人間の動作、発話、状況画像の情報が入力される。（2）ロボットの動作、発話の情報が出力される。（3）画像に関する質問発話と動作指示発話などの複数の種類の発話行為に対して適応的に応答可能である。最初の二年間で開発した手法のように格助詞を含む文とその意味を学習するだけでなく、マルチモーダルな応答までも End-to-end で学習することに成功したことは、計画当初の予想を上回る極めて大きな成果であると言える。

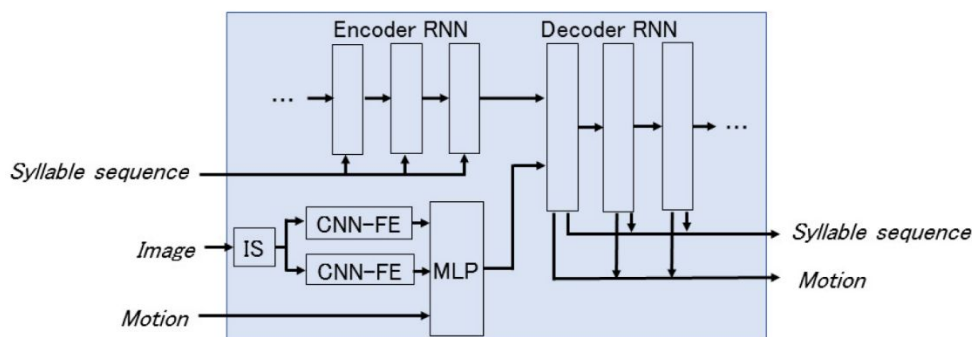


図 2 マルチモーダル・コミュニケーション・ニューラルネットワークモデル

最終年度（30年度）は、人間の認知メカニズムの構成論的な理解の研究として、MCNNM を利用し、統合失調症の計算モデルの構築に成功した。健常者の対話を表すデータを用いて学習したネットワークの一部に変形（障害）を加えることで、統合失調症の患者と同様な発話応答を出力することを確認した。

以上に述べたように、本研究採択時に計画していたすべて研究項目について、実現方法は異なるものの、想定を大きく上回る成果を上げることに成功した。

5 . 主な発表論文等

〔雑誌論文〕(計 2 件)

Tadahiro Taniguchi, Takayuki Nagai, Tomoaki Nakamura, Naoto Iwahashi, Tetsuya Ogata, Hideki Asoh, Symbol Emergence Robotics: A Survey, *Advanced Robotics*, 査読有, Vol.30, 2016, pp. 706-728

T. Taniguchi, D. Mochihashi, T. Nagai, S. Uchida, N. Inoue, I. Kobayashi, T. Nakamura, Y. Hagiwara, N. Iwahashi, T. Inamura, A Survey on Frontiers of Language and Robotics, *Advanced Robotics*, 査読有, 2019 (to appear)

〔学会発表〕(計 3 件)

Takabuchi, N. Iwahashi, Takeo Kunishima, A Language Acquisition Method Based on Statistical Machine Translation for Application to Robots, *The Sixth Joint IEEE Int. Conf. Developmental Learning and Epigenetic Robotics*, 2016.

K. Ye, K. Takabuchi, N. Iwahashi, DNN-based Language Acquisition for Object Manipulation Task, *The Seventh Joint IEEE Int. Conf. Developmental Learning and Epigenetic Robotics*, 2017.

守屋綾祐, 高瀬健太, 岩橋直人, Neural Network Model for Human-Robot Multimodal Linguistic Interaction, *HAI シンポジウム*, 2018.

〔その他〕
ホームページ等
<http://www-ail.c.oka-pu.ac.jp/>

6 . 研究組織

研究代表者のみである。

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属されます。