

科学研究費助成事業 研究成果報告書

平成 29 年 6 月 12 日現在

機関番号：12611

研究種目：基盤研究(B) (一般)

研究期間：2014～2016

課題番号：26280096

研究課題名(和文) 潜在意味空間において感覚情報を言語化し言語的思考を行うロボットの実現

研究課題名(英文) Development of a robot which can verbalize its sensing information in the latent semantic space and handles the information with language

研究代表者

小林 一郎 (Kobayashi, Ichiro)

お茶の水女子大学・基幹研究院・教授

研究者番号：60281440

交付決定額(研究期間全体)：(直接経費) 12,500,000円

研究成果の概要(和文)：本申請研究では、(i) マルチモーダル情報を言語化するための語彙概念・文法概念の取得、(ii) 言語資源の転移学習、(iii) 言葉でコミュニケーションをしてロボットを動作させる言語の意味に基づくロボットの動作生成、(iv) 統語情報を考慮したテキスト生成手法の開発、(v) 強化学習を用いたロボットによる言葉の概念の獲得について開発を進めた。

研究成果の概要(英文)：We have worked on five topics: (i) Obtaining the concepts of words and grammar for verbalizing multimodal information (ii) Transfer learning of linguistic resources (iii) Action generation of a robot based on the meaning of language for making a robot act through communicating with language, (iv) Development of natural language generation with Monte Carlo tree search, and (v) Language acquisition for a robot by using reinforcement learning with multimodal information.

研究分野：知能情報処理、言語情報処理

キーワード：マルチモーダル情報 言語生成 動作生成 潜在的意味 時系列データ ロボット モンテカルロ木探索 概念獲得

1. 研究開始当初の背景

高齢化社会へ移行する中で日常生活を支援するロボットとの円滑なコミュニケーションが望まれている。しかし、現時点におけるコミュニケーションは、十分に円滑なものとはなっていない。我々は、その主要な原因の一つが、ロボットが観測する状況を上手く記号化(言語化)することができていないため、人との状況認識の間に齟齬が生じていることにあると考えている。ロボットが観測する状況は、ロボットが持つ各種センサから観測する視覚、聴覚、感触などのマルチモーダルな感覚情報によって形成され、それらは高次元の複数の時系列データとして観測される。ロボットが人間と同じように状況を認識し、円滑なコミュニケーションを実現するためには、それらの感覚時系列データを適切なレベルで言語化する手法、および言語化によって認識された結果を適切に人の発話と結び付ける手法の開発が必要である。

非言語情報の言語化は世界的にも重要な課題と認識され始めている。たとえば、2013年の自然言語処理分野で最難関とされるACLという国際会議において、最優秀論文賞に動画中の人の動作を説明する文生成を行う研究 [Yu+ 2013] が選ばれ、また、同会議にて動画に映る調理の動作を説明する文生成の研究 [Regneri+ 2013] の報告や「言語処理へ向けた画像処理の基礎」に関するチュートリアルが開かれるなどしている。しかし、彼らを含む多くの研究は、視覚情報一つの時系列データからの文生成に関する研究であり、ロボットの研究分野で扱われているようなマルチモーダル情報を解釈統合することより一つの現象を説明する文の生成は行っていない。研究代表者の小林は、平成 19 - 21 年度および平成 23 - 25 年度において科学研究費補助金基盤(C)にて時系列データの言語化の基盤技術の開発を行い、複数のドメインにおける時系列データの言語化、及び、複数の時系列データ間の関係を説明する言語化においてその有効性を検証し、その成果を基に人の関節の動きを捉えた高次元の時系列データの言語化に取り組んでいる。研究分担者の長井は、平成 20 - 22 年度および平成 23 - 25 年度において科学研究費補助金(C)にて、マルチモーダル情報に基づくロボットの現実世界理解に関する研究を行い、ロボットの感覚情報から物体認識と人の行動の統合的概念の形成に取り組んでいる。

これまでの小林の研究は、時系列データの現象を説明するのに尤もらしい文を生成することに成功しているが、語彙の概念を獲得するなどの学習による言語化の視点が欠如している。一方、現時点での長井の研究は、マルチモーダル情報と言語表現からの学習によって語彙概念の獲得をすることに成功しているが、単語レベルにとどまっておらず、語彙概念を組み合わせた文章を扱うには至っていない。

2. 研究の目的

これまでの研究成果を補完的に組み合わせることでさらに発展させることにより、ロボットの感覚情報として得られる複数の時系列データから、物体と人の行動の統合概念を文として生成可能にする新しい言語化手法を開発する。それを実現するために、以下の5つの課題について研究を進展させる。

- (i) マルチモーダル情報を言語化するための語彙概念・文法概念の取得
- (ii) 言語資源の転移学習
- (iii) 言語の意味に基づくロボットの動作生成
- (iv) 統語情報を考慮したテキスト生成手法の開発
- (v) 強化学習を用いたロボットによる言葉の概念の獲得について開発

3. 研究の方法

- (i) マルチモーダル情報を言語化するための語彙概念・文法概念の取得
実環境においてロボットが観測する情報はマルチモーダルな情報である。語意獲得は、ロボットが経験することによって取得するマルチモーダルな知覚情報のカテゴリ分類を基盤としており、Latent Dirichlet Allocation (LDA) をマルチモーダル情報の分類に適用したMultimodal LDA (MLDA) [Nakamura 09]を用いて行う。mMLDA を用いて階層的に概念を形成することが可能になる。同時にユーザが教示文を与えることで単語と形成された概念との結びつきを相互情報量を基準にして学習する。しかし、概念との相互情報量だけでは正確な単語と概念の結びつきを見つけることは難しい。そのため、Bayesian HMMを用いて統語情報を用いた文法の学習を実現する。
[詳細は業績 1 (雑誌), 1,7,9,18 を参照]

- (ii) 言語資源の転移学習
ロボットが持つセンサから得られる時系列データを説明する自然言語文から作られたバイグラム言語モデルを用いて、最適な単語の組み合わせを発見することより文を生成する。この際、言語資源(バイグラムモデル)が存在しない場合についてもテキスト生成を可能にするように動作の意味体系を考慮して欠損している言語資源を推定する最小二乗推定を用いた Zero-shot 学習を提案する。
[詳細は業績 3,10,14,20,29,38 を参照]

- (iii) 言語の意味に基づくロボットの動作生成
word2vec を用いて語彙の分散意味表現ベクトルとロボットの動作ベクトルとの対応関係を学習し、入力として与えら

れた新しい語彙の意味を反映するロボットの動作生成（調理動作を対象とする）を実現し、「切る」「まぜる」などの調理の基本動作に対して、副詞の意味がかかわる動作生成を実現する。
[詳細は業績 5,13,17,21,28,34,37 参照]

- (iv) 統語情報を考慮したテキスト生成手法の開発
モンテカルロ木探索 (MCTS) のアルゴリズムを用いて、適切な統語構造および語彙選択をシミュレーションにおいて探索し、テキスト生成を行う手法を開発する。統語構造に文脈自由文法を用い、MCTS のシミュレーションにおける一つのノードが成長する構文木となり、シミュレーションの最終局面で生成された文を評価（正文か非文）することにより、文生成を行う手法を提案した。
[詳細は業績 2,11,15,22,30,33,36 を参照]
- (v) 強化学習を用いたロボットによる言葉の概念の獲得について開発
ロボットが観測するマルチモーダル情報から動作の概念を mMLDA を用いて取得し、強化学習をする際に概念を取得するたびに状態空間を獲得した概念で表現し直し、探索する空間の領域を狭め、学習の効率性を実現する手法を提案した。
[詳細は業績 4,12,16,19,23,27,35 を参照]

4. 研究成果

課題(i)においては、マルチモーダル情報から教師なしで語彙、文法を獲得することに成功した。論文誌 1 本、トップクラスの国際会議 1 本、国内発表 4 本の成果報告を行った。

課題(ii)においては、限られた動作の意味体系ではあるが、最小二乗推定を用いた zero-shot 学習がうまく機能することが確認できた。国際会議 1 本、国内発表 5 本の成果報告を行った。

課題(iii)においては、ロボットの動作に程度を表す言葉の意味を反映することに成功した。国際会議 1 本、国内発表 6 本の成果報告を行った。

課題(iv)においては、MCTS を用いて、統語構造を正しく構築する文生成の手法を確立できた。国際会議 1 本、国内発表 6 本の成果報告を行った。

課題(v)においては、マルチモーダル情報を用いたロボットの行動の学習において階層型マルチモーダル LDA を利用することによりマルチモーダル情報から概念を獲得し、獲得した概念に基づいてさらに学習を効率良く行う手法の提案を行った。国際会議 1 本、国内発表 6 本の成果報告を行った。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 1 件)

1. Muhammad Attamimi, Yuji Ando, Tomoaki Nakamura, Takayuki Nagai, Daichi Mochihashi, Ichiro Kobayashi, and Hideki Asoh: Learning word Meaning and Grammar for Verbalization of Daily Life Activities Using Multilayered Multimodal Latent Dirichlet Allocation and Bayesian Hidden Markov Models, *Advanced Robotics*, DOI:10.1080/01691864.2016.1172507, 2016.

[学会発表](計 38 件)

1. 熊谷香織, 小林一郎, 持橋大地, 麻生英樹, 中村友昭, 長井隆行, モンテカルロ木探索を用いた構文木構築に基づく非決定的な文生成, 6Q-05, 情報処理学会第 79 回全国大会, 名古屋大学, 2017 年.
2. 恒川英里, 小林一郎, 麻生英樹, 持橋大地, 中村友昭, 長井隆行, 強化学習を用いたロボットの行動に関する概念獲得への取り組み, 3M-06, 情報処理学会第 79 回全国大会, 名古屋大学, 2017 年.
3. 濱園侑美, 小林一郎, 持橋大地, 麻生英樹, 中村友昭, 長井隆行, 副詞の意味に基づくロボットの動作生成に関する一考察, 7Q-06, 情報処理学会第 79 回全国大会, 名古屋大学, 2017 年.
4. 樺山絵里, 小林一郎, 持橋大地, 麻生英樹, 中村友昭, 長井隆行, zero-shot 学習を用いた言語資源の転移に関する一考察, 7Q-07, 情報処理学会第 79 回全国大会, 名古屋大学, 2017 年.
5. 熊谷香織, 小林一郎, 持橋大地, 麻生英樹, 中村友昭, 長井隆行, モンテカルロ木探索を用いた構文木構築に基づく頑健な文生成, P1-4, 言語処理学会第 23 回年次大会, つくば, 2017 年.
6. 濱園侑美, 小林一郎, 麻生英樹, 中村友昭, 長井隆行, 持橋大地, 曖昧言語表現に対するロボット動作の対応関係学習, P6-3, 言語処理学会第 23 回年次大会, つくば, 2017 年.
7. Eri Kabayama, Muhammmad Attamimi, Ichiro Kobayashi, Hideki Asoh, Daichi Mochihashi, Tomoaki Nakamura, and Takayuki Nagai: Evaluation of the Sentences Generated Based on Language Model Applied by Zero-shot Learning, Workshop on Machine Learning Methods for High-Level Cognitive Capabilities in Robotics 2016, Daegeon, Korea, Oct.2016.
8. Eri Tsunekawa, Muhammmad Attamimi, Ichiro Kobayashi, Hideki Asoh, Daichi

- Mochihashi, Tomoaki Nakamura, and Takayuki Nagai: An Approach to Making a Plan for Tidying up with the Action Concepts Acquired by Multimodal Information, Workshop on Machine Learning Methods for High-Level Cognitive Capabilities in Robotics 2016, Daegeon, Korea, Oct.2016.
9. Yumi Hamazono, Ichiro Kobayashi, Hideki Asoh, Daichi Mochihashi, Muhammad Attamimi, Tomoaki Nakamura, and Takayuki Nagai: Learning the Correspondence between Distributed Semantics of Words and Robot 's Action, Workshop on Machine Learning Methods for High-Level Cognitive Capabilities in Robotics 2016, Daegeon, Korea, Oct.2016.
 10. Tomoaki Nakamura, Kensuke Iwata, Takayuki Nagai, Daichi Mochihashi, Ichiro Kobayashi, Hideki Asoh and Masahide Kaneko: Continuous Motion Segmentation Based on Reference Point Dependent GP-HSMM, IROS2016: Workshop on Machine Learning Methods for High-Level Cognitive Capabilities in Robotics, Oct. 2016
 11. Kaori Kumagai, Ichiro Kobayashi, Daichi Mochihashi, Hideki Asoh, Muhammad Attamimi, Tomoaki Nakamura, and Takayuki Nagai: Human-like Natural Language Generation Using Monte Carlo Tree Search, Workshop on Computational Creativity in Natural Language Generation(CC-NLG), Edinburgh, Sep.5-6,2016.
 12. 岩田健輔, 中村友昭, 長井隆行, 持橋大地, 小林一郎, 麻生英樹: 参照点に依存したガウス過程隠れセミマルコフモデルに基づく連続動作の分節化, 日本ロボット学会学術講演会, 3Z2-07, Sep. 2016.
 13. 中村友昭, Attamimi Muhammad, 長井隆行, 持橋大地, 小林一郎, 麻生英樹, 金子正秀, ガウス過程の隠れセミマルコフモデルに基づく身体動作の分節化, 103-5, 第30回人工知能学会全国大会, 小倉, 6月, 2016.
 14. 恒川英里, Attamimi Muhammad, 小林一郎, 中村友昭, 長井隆行, 麻生英樹, 持橋大地, 多層マルチモーダルLDAと強化学習による片付け概念獲得への取り組み, 103-4in1, 第30回人工知能学会全国大会, 小倉, 6月, 2016.
 15. 濱園侑美 小林一郎 麻生英樹 持橋大地 Muhammad Attamimi 中村友昭 長井隆行, 語彙の分散意味表現とロボット動作との対応関係の学習, 105-0S-22b-3in1, 第30回人工知能学会全国大会, 小倉, 6月, 2016.
 16. 樺山 絵里, 小林一郎, 麻生英樹, 持橋大地, Attamimi Muhammad, 中村友昭, 長井隆行, 人の挙動を表現するテキスト生成の一考察, 3P1-2in2, 第30回人工知能学会全国大会, 小倉, 6月, 2016.
 17. 熊谷香織, 持橋大地, 小林一郎, 麻生英樹, Attamimi Muhammad, 中村友昭, 長井隆行, モンテカルロ木探索を用いた構造的正しさと言語モデルを考慮した文生成, 304-0S-04b-5in1, 第30回人工知能学会全国大会, 小倉, 6月, 2016.
 18. 安東裕司 Attamimi Muhammad 中村友昭 長井隆行, 持橋大地, 小林一郎, 麻生英樹, 多層マルチモーダルLDAと隠れセミマルコフモデルを用いた概念・語彙・文法の相互学習, 304-0S-04b-4, 第30回人工知能学会全国大会, 小倉, 6月, 2016.
 19. 樺山絵里, 麻生英樹, 小林一郎, 持橋大地, Muhammad Attamimi, 中村友昭, 長井隆行: 言語モデルを用いた人の動作を説明するテキスト生成およびその評価, P1-5, 第22回言語処理学会年次大会, 東北大学, 3月7-11日, 2016.
 20. 濱園侑美, 小林一郎, 麻生英樹, 持橋大地, Muhammad Attamimi, 中村友昭, 長井隆行: 語彙の分散意味に基づくヒューマノイドロボットの動作生成に向けた基礎的検討, P1-4, 第22回言語処理学会年次大会, 東北大学, 3月7-11日, 2016.
 21. 熊谷香織, 持橋大地, 小林一郎, 麻生英樹, Muhammad Attamimi, 中村友昭, 長井隆行: モンテカルロ木探索を用いた統語情報を考慮した文生成, P9-3, 第22回言語処理学会年次大会, 東北大学, 3月7-11日, 2016.
 22. 恒川英里, Attamimi Muhammad, 小林一郎, 長井隆行, 中村友昭, 麻生英樹, 持橋大地: 強化学習を用いた物体の片付けに関する概念獲得への取り組み, (学生奨励賞受賞), 情報処理学会第78回全国大会講演論文集, 3P-07, pp. 2-359-360, 2016.
 23. アッタミムハンマド, 安東裕司, 中村友昭, 長井隆行, 持橋大地, 小林一郎, 麻生秀樹, 人の活動の言語化のための語意と文法の学習, 計測自動制御学会システム・情報部門学術講演会 2015, 函館, 11月, 2015.
 24. 中村友昭, アッタミムハンマド, 長井隆行, 持橋大地, 小林一郎, 麻生英樹, 金子正秀, ガウス過程の隠れマルコフモデルによる時系列データ分節化, 計測自動制御学会システム・情報部門学術講演会 2015, 函館, 11月, 2015.
 25. Muhammad Attamimi, Yuji Ando, Tomoaki Nakamura, Takayuki Nagai, Daichi Mochihashi, Ichiro Kobayashi, Hideki Asoh: Learning Word Meanings and Grammar for Describing Everyday Activities in Smart Environments.

- EMNLP 2015: 2249-2254.
26. 樺山 絵里、麻生 英樹、Attamimi Muhammad、小林 一郎、持橋 大地、中村友昭、長井 隆行、現象の意味的構造に基づく言語モデルの zero-shot 学習の試み、2F3-2、第 29 回人工知能学会全国大会、はこだて未来大学、5 月、2015.
 27. 熊谷 香織、持橋 大地、小林 一郎、麻生 英樹、Attamimi Muhammad、中村友昭、長井 隆行、モンテカルロ木探索を用いた確率文脈自由文法に基づくテキスト生成、2F3-3、第 29 回人工知能学会全国大会、はこだて未来大学、5 月、2015.
 28. 恒川 英里、小林 一郎、麻生 英樹、持橋 大地、Attamimi Muhammad、中村友昭、長井 隆行、ヒューマノイドロボットの画像認識に基づく行動知識の獲得への取り組み、2D4-1、第 29 回人工知能学会全国大会、はこだて未来大学、5 月、2015.
 29. 濱園 侑美、小林 一郎、麻生 英樹、持橋 大地、Attamimi Muhammad、中村友昭、長井 隆行、ヒューマノイドロボットの動作生成に向けた基礎的検討、2D1-0S-12a-1、第 29 回人工知能学会全国大会、はこだて未来大学、5 月、2015.
 30. 安東 裕司、Muhammad Attamimi、中村 友昭、長井 隆行、持橋 大地、小林 一郎、麻生 英樹、多層マルチモーダル LDA と HMM を用いた文法の学習、2F4-0S-01a-6、第 29 回人工知能学会全国大会、はこだて未来大学、5 月、2015.
 31. 長井 隆行、中村 友昭、Attamimi Muhammad、持橋 大地、小林 一郎、麻生 英樹、多層マルチモーダル LDA と強化学習による意味理解に基づく行動決定、2F4-0S-01a-7、第 29 回人工知能学会全国大会、はこだて未来大学、5 月、2015.
 32. 熊谷香織、持橋 大地、小林 一郎、麻生 英樹、Muhammad Attamimi、中村友昭、長井 隆行、モンテカルロ木探索を用いた確率文脈自由文法に基づくテキスト生成、第 21 回言語処理学会年次大会、京都大学、3 月、2015 .
 33. 恒川英里、小林 一郎、麻生 英樹、長井 隆行、中村友昭、持橋 大地、画像認識に基づくロボットの行動を制御する強化学習の取り組み、第 77 回情報処理学会全国大会、京都大学、3 月、2015 .(学生奨励賞受賞)
 34. 濱園侑美、小林 一郎、麻生 英樹、長井 隆行、中村友昭、持橋 大地、ロボットの調理動作学習へ向けた取り組み、第 77 回情報処理学会全国大会、京都大学、3 月、2015 .
 35. 樺山絵里、麻生 英樹、小林 一郎、持橋 大地、Muhammad Attamimi、中村友昭、長井 隆行、Zero-shot 学習した言語モデルによるテキスト生成結果の評価、第 21 回言語処理学会年次大会、京都大学、3 月、2015 .

36. 安東裕司、アッタミミ・ムハンマド、中村友昭、長井 隆行、持橋 大地、小林 一郎、麻生 英樹、日常生活言語化のためのデータ取得システム、計測自動制御学会、12 月、2014 .
37. Muhammad Attamimi、中村友昭、長井 隆行、持橋 大地、小林 一郎、麻生 英樹、獲得した概念階層と語彙に基づく文生成、第 32 回日本ロボット学会学術講演会、九州産業大学、9 月 4 日-6 日、2014 .
38. 麻生 英樹、小林 一郎、視覚情報と言語情報の対応づけにおける言語モデルの転移学習、日本行動計量学会第 42 回大会、東北大学、9 月 2 日-4 日、2014 .

〔図書〕(計 0 件)

〔産業財産権〕

出願状況(計 0 件)

名称：
 発明者：
 権利者：
 種類：
 番号：
 出願年月日：
 国内外の別：

取得状況(計 0 件)

名称：
 発明者：
 権利者：
 種類：
 番号：
 取得年月日：
 国内外の別：

〔その他〕
 ホームページ等

6. 研究組織

- (1) 研究代表者
 小林一郎 (KOBAYASHI, Ichiro)
 お茶の水女子大学・基幹研究院・教授
 研究者番号：60281440
- (2) 研究分担者
 長井隆行 (NAGAI, Takayuki)
 電気通信大学大学院・情報理工学研究所・
 機械知能システム工学専攻・教授
 研究者番号：40303010
- (3) 研究分担者
 持橋大地 (MOCHIHASHI, Daichi)
 統計数理研究所・数理・推論研究系・准教授

研究者番号：80418508

(4) 連携研究者

麻生英樹 (ASOH, Hideki)

産業技術総合研究所・人工知能研究センタ

ー・副研究センター長

研究者番号：10344194