

科学研究費助成事業 研究成果報告書

平成 30 年 6 月 12 日現在

機関番号：34315

研究種目：若手研究(B)

研究期間：2016～2017

課題番号：16K16133

研究課題名(和文)クラウド仮想環境に基づく社会的マルチモーダル概念学習基盤

研究課題名(英文)Platform for concept formation based on multimodal information using a cloud VR system

研究代表者

萩原 良信 (Yoshinobu, Hagiwara)

立命館大学・情報理工学部・助教

研究者番号：20609416

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：本研究では、大規模な人とロボットのインタラクションによるマルチモーダル情報に基づく物体・場所概念の学習を可能とするクラウド仮想環境：SIGVerse+Lを構築した。まず、被験者の音声、視線、身体動作を仮想環境のアバターに反映する機能を構築した。次に、仮想環境のロボットが被験者の音声、環境の画像、地図上の位置などのマルチモーダル情報を収集する機能をROSにより実装した。最後に、SIGVerse+Lを用いたヒューマンロボットインタラクションに基づく場所概念学習をマルチモーダル情報に基づくベイズ生成モデルにより達成した。

研究成果の概要(英文)：In this research, we have constructed SIGVerse + L, a cloud virtual environment that enables learning of object and location concepts based on multimodal information by interaction between large-scale people and robots. First, we constructed a function to reflect the subject's voice, line of sight, and body motion on the avatar of the virtual environment. Next, the ROS implements the function of collecting multimodal information such as speech of the subject, image of the environment, position on the map, etc. by the robot in the virtual environment. Finally, location concept learning based on human robot interaction using SIGVerse + L was achieved by Bayesian model based on multimodal information.

研究分野：知能ロボティクス

キーワード：マルチモーダル 概念 サイバーフィジカル ベイズモデル ヒューマンロボットインタラクション

1. 研究開始当初の背景

家庭環境において人と共存するロボットは、人間とのやり取りのパターン、物体や場所の名前やその意味を構成する概念など、様々な知識を得る必要がある。

電気通信大学の長井らは、ロボットの視覚、触覚、聴覚情報から構成される Multimodal LDA に基づいて物体の概念を自律的に学習するボトムアップアプローチにより、人と類似する物体のカテゴリ分類を実現した[1]。また、立命館大学の谷口らは、Multimodal LDA の拡張により、人とロボットの対話に基づく場所概念と語彙獲得の相互学習を実現している[2]。これらの研究では、画像や音声、触覚情報などで構成されるマルチモーダル情報に基づいて物体や場所の概念をロボットがボトムアップに獲得している。

しかしながら、一人の被験者から 125 個の物体を学習するのに一週間程度、1 部屋の場所を学習するのに数日を要している。物体や場所の概念を環境や状況に対して汎化するには多様な条件下での大規模な学習が必須であり、これらの学習をより短期間で効率的に実施する仕組みが必要である。

2. 研究の目的

本研究では、環境や状況に対して汎化された物体や場所の概念の学習を可能とするため、多様な環境における複数のユーザからのマルチモーダル情報に基づく概念学習を仮想世界において効率的に実施する基盤技術を創生する。具体的には、図 1 のように、没入型インタフェースを用いて仮想世界にログインした複数のユーザから、物体や場所についての概念をマルチモーダル情報に基づいて学習し、その知識統合による概念の形成を可能とするクラウド仮想環境 (SIGVerse+L) を構築する。これは、仮想世界におけるマルチモーダル情報の収集を可能とする没入型インタフェースと対話機能、認識機能を社会的知能発生学シミュレータ (SIGVerse) [3] に実装したクラウド仮想環境である。さらに、その知識をロボットミドルウェアである ROS (Robot Operating System) を通じて実世界のロボットの知識として転移する仕組み、理論、手法を確立する。

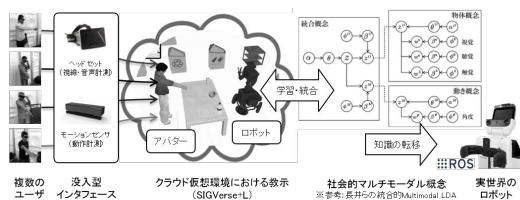


図 1. SIGVerse+L の概要図

3. 研究の方法

(1) SIGVerse+L の構築

マルチモーダル情報に基づく大規模な概念の学習を可能とする没入型インタフェースを伴う

クラウド仮想環境 (SIGVerse+L) を構築する。具体的には、実世界の人間の音声、視線、全身動作を高精度に仮想世界のアバターに反映する最新の没入型インタフェースと、マルチモーダル情報に基づく人からの概念学習を可能とする対話機能の SIGVerse への適切な実装を行う。人間の音声、視線、全身動作をアバターに反映するデバイスとして、マイク付属の 3D ヘッドセット、全身モーションセンサを使用する予定である。さらに、人からの概念学習で重要となる音声対話機能を強化するため、NICT の杉浦らの開発した音声認識・合成エンジン (rospeex) [4] を SIGVerse に実装する。

(2) マルチモーダル概念の形成

構築したクラウド仮想環境 (SIGVerse+L) を用いて、数人のユーザからの教示によるマルチモーダル情報に基づく概念形成を実施する。異なるユーザからの教示によるマルチモーダル概念の知識統合は、長井らの統合的 Multimodal LDA を発展させることにより実現する[5]。研究代表者は、実機ロボットの階層的な場所概念の獲得に関する研究[6]において既に Multimodal LDA を用いており、この応用によりマルチモーダル概念の形成を実現する。

(3) 知識転移インタフェースの構築

SIGVerse+L で学習したマルチモーダル概念の知識を実空間のロボットに転移するインタフェースを構築する。具体的には、統合的 Multimodal LDA に基づくマルチモーダル概念学習を ROS などのロボット共通のミドルウェアのノードとして実装し、このノードを実世界のロボットに転移する事によって実現する。このとき、学習したモデルパラメータの転移を可能とする計算論モデルを構築する。

(4) 有用性の検証実験

マルチモーダル概念学習における SIGVerse+L の有用性を検証する。まず、研究代表者が以前に所属していた国立情報学研究所と現在所属している立命館大学に実験環境を構築し、マルチモーダル概念学習の事前実験を実施する。この事前実験における問題点を修正した後、本研究で構築したマルチモーダル概念学習に基づくクラウド仮想環境 (SIGVerse+L) 及び、仮想世界で学習した知識を実世界のロボットに転移するインタフェースをオープンソースで公開する。公開した SIGVerse+L と知識転移インタフェースを用いて、知能ロボットの世界的な競技会である RoboCup 実行委員会の協力を得ながら社会的マルチモーダル概念学習における SIGVerse+L の検証実験及び、社会実験を実施し、その有用性を検証する。

4. 研究成果

(1) SIGVerse+L の構築

図 2 に構築した SIGVerse+L の概要を示す没入型インタフェース HTC Vive (HTC 社製) を用いてユーザのアバターへの没入と身体を制御を実現した。さらに、ユーザの音声を認識して言語情報に変換する機能として、多言語辞書が使用可能な rospeek と言語モデルの構築が容易な Julius[7] を ROS により実装した。さらに、仮想空間での自己位置を SLAM[8] により認識する機能、仮想環境の撮像画像から CNN (Convolutional Neural Network) のフレームワークである Caffe[9] により特徴を抽出する機能を ROS により実装した。さらに、人に追従しながら物体や場所の概念の学習を可能とする機能も ROS により構築した。これにより、仮想環境における人とロボットのインタラクションから音声、画像、位置のマルチモーダル情報を収集するシステムを構築した。

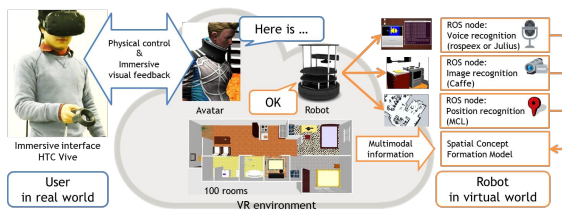


図 2. 構築したシステムの概要図

(2) 社会的マルチモーダル概念の形成

SIGVerse+L を用いて、数人のユーザからの教示に基づくマルチモーダル場所概念の形成を実施した。場所概念の形成には、マルチモーダル情報を観測としたベイズ生成モデルを用いた。具体的には、画像と言語情報のクラスタリングを行う Multimodal LDA と位置情報のクラスタリングを行う混合ガウス分布、自己位置を推定する MCL (Monte Carlo Localization) から構成される場所概念形成モデル[10]を用いた。これにより、場所概念学習及び評価実験を仮想環境で可能とした。

実際に、仮想環境においてユーザからの場所の単語 (Word) の発話とマルチモーダル情報から場所概念学習を行い、学習結果に基づく場所の名前の予測の評価実験を行った。図 3 が評価実験の結果である。横軸がデータに対するユーザからの場所の名前の発話の割合、縦軸が各割合において学習した場所概念に基づく場所の名前予測の正解率である。ユーザからの名前の教示が増えるに従って場所の学習が進み正解率が向上している事が確認できる。

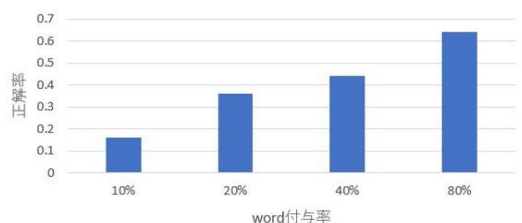


図 3. 場所概念学習における名前予測の評価

(3) 知識転移のインタフェースの構築

SIGVerse+L で学習したマルチモーダル概念の知識を実空間のロボットに転移するインタフェースを構築した。知識転移のインタフェースは、SIGVerse と実世界のロボットで共通に利用されるミドルウェアである ROS によって構築した。具体的には、仮想空間で得られた視覚、音声、位置情報を ROS のノードを通じて CNN による画像認識、rospeek による音声認識、MCL による位置認識と接続し、これらのマルチモーダル情報を場所概念の転移学習モデルの観測として与え、Gibbs sampling による潜在変数の推定を行う事で仮想世界と実世界の間で転移可能な概念を構築した。具体的には、仮想環境で学習された場所概念形成モデルのモデルパラメータを実環境の場所概念学習に転移する転移学習モデルを図 4 のように構築した。このモデルを用いて実環境間、仮想環境間での場所概念の転移学習を達成した。今後、実環境と仮想環境の間での転移学習実験に取り組む予定である。

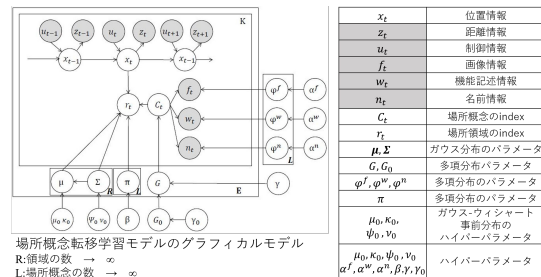


図 4. 場所概念の転移学習モデル

(4) 有用性の検証実験

マルチモーダル概念学習における SIGVerse+L の有用性を検証した。まず、複数の被験者と 20 種類の家庭環境を用いたヒューマンロボットインタラクションに基づく場所概念学習実験を実施し、環境に対して汎化された場所概念の形成実験を実施した。図 5 に実験の様子を示す。実験では、図 5 左のような仮想の家庭環境において被験者が仮想環境のロボットへの場所の教示と学習を行った。右はユーザの様子と音声対話におけるテキストである。実験環境におけるヒューマンロボットインタラクションに基づく学習は実世界では高コストであるが、SIGVerse+L により実験環境の構築、ロボットの維持管理のコストを大幅に低減する事ができた。



図 5. 評価実験の様子

(5) その他の研究成果

2017年5月に実施された RoboCup Japan Open @Home Simulation League の Finals , 2018年5月に実施された経済産業省主催の World Robot Summit Service Robot Challenge プレイベントにおいて当該研究課題のデモンストレーションを行い, SIGVerse+L の有用性を国内外のロボティクス研究者に示した. これらの取り組みを通じて, 国立情報学研究所, 岡山県立大学をはじめとする複数の教育研究機関との共同研究が開始され, 本研究課題はさらなる広がりを見せている. 当該研究課題に関するデモンストレーションによる知能ロボットの競技会における受賞は以下である.

RoboCup Japan Open 2018 @Home Domestic Standard Platform League, 三位入賞, OIT Challenger & Duckers, 2018年5月, 岐阜.

RoboCup Japan Open 2018 @Home Education League, 三位入賞, Ri-one & Duckers, 2018年5月, 岐阜.

World Robot Summit Partner Robot Challenge (Virtual Space) Pre-event, 三位入賞, Duckers, 2018年5月, 岐阜.

Team「Duckers」, RoboCup Japan Open 2017 @Home Simulation League 三位入賞, 2017年5月, 愛知.

Team「Ri-one & Duckers」, RoboCup Japan Open 2017 @Home Education League 三位入賞, 2017年5月, 愛知.

また, 本研究で開発した場所概念形成モデル及び, 実験用の家庭環境(100種類), 各種の ROS ノードは, 研究代表者が管理する RoboCup@Home のチーム Duckers のウェブサイト (<https://emlab.jimdo.com/multimedia/>) からダウンロード可能である. また, 場所概念に関する以下のソースコードを GitHub で公開している.

Spatial Concept Formation, マルチモーダル情報からベイズ生成モデルにより場所概念を形成するソースコード, URL: https://github.com/EmergentSystemLabStudent/Spatial_Concept_Formation
Spatial Concept based Semantic Mapping, 場所概念に基づいてセマンティックマップを生成するソースコード, URL: https://github.com/EmergentSystemLabStudent/spco_mapping

<引用文献>

- [1] T. Nakamura, N. Iwahashi, T. Nagai, Multimodal categorization by hierarchical dirichlet process, In Proceedings of 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems,

- pp.1520–1525, 2011.
- [2] A. Taniguchi, T. Taniguchi, and T. Inamura, Spatial concept acquisition for a mobile robot that integrates self-localization and unsupervised word discovery from spoken sentences, IEEE Transactions on Cognitive and Developmental Systems, Vol.8, pp.285–297, 2016.
- [3] T. Inamura et al, Simulator platform that enables social interaction simulation -SIGVerse: SocioIntelliGenesis simulator-, in IEEE/SICE International Symposium on System Integration, pp.212–217, 2010.
- [4] K. Sugiura, and K. Zettsu, Rospeex: A cloud robotics platform for human-robot spoken dialogues, In Proceedings of 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp.6155–6160, 2015.
- [5] アッタミ, ムハンマド, 阿部, 中村, 船越, 長井, 多層マルチモーダル LDA を用いた人の動きと物体の統合概念の形成, 日本ロボット学会誌, Vol.32, no.8, pp89-100, 2014.
- [6] Y. Hagiwara, M. Inoue, and T. Taniguchi, Place concept learning by hmlda based on position and vision information, In 13th IFAC / IFIP / IFORS / IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems, Vol. 49, pp.216–220, 2016.
- [7] 李晃伸, 大語彙連続音声認識エンジン Julius ver. 4, 情報処理学会研究報告音声言語情報処理 (SLP), Vol.129, pp.307-312, 2007.
- [8] H. Durrant-Whyte, and T. Bailey, Simultaneous localization and mapping: part i. IEEE robotics & automation magazine, Vol.13, pp.99–110, 2006.
- [9] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, et al. Caffe: Convolutional architecture for fast feature embedding. In Proceedings of the 22nd ACM international conference on Multimedia (ACM), pp.675–678, 2014.
- [10] S. Ishibushi, A. Taniguchi, T. Takano, Y. Hagiwara, and T. Taniguchi, Statistical localization exploiting convolutional neural network for an autonomous vehicle. In 41th Annual Conference of the IEEE Industrial Electronics Society, pp.1369–1375, 2015.

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 3 件)

Hagiwara Yoshinobu, Inoue Masakazu,

Kobayashi Hiroyoshi, Taniguchi Tadahiro, Hierarchical Spatial Concept Formation Based on Multimodal Information for Human Support Robots, *Frontiers in Neurorobotics*, 査読有, Vol.12, Article 11, 2018, pp.1-16

DOI: 10.3389/fnbot.2018.00011

Isobe Shota, Taniguchi Akira, Hagiwara Yoshinobu, Taniguchi Tadahiro, Learning Relationships Between Objects and Places by Multimodal Spatial Concept with Bag of Objects, *Lecture Notes in Computer Science*, Springer, Cham, 査読有, Vol.10652, 2017, pp.115-125 [Best Interactive Session Finalist]

DOI: 10.1007/978-3-319-70022-9_12

Taniguchi Akira, WanPeng Lv, Taniguchi Tadahiro, Takano Toshiaki, Hagiwara Yoshinobu, Yano Shiro, Simultaneous Localization, Mapping and Self-body Shape Estimation by a Mobile Robot, *Advances in Intelligent Systems and Computing*, Springer, Cham, 査読有, Vol.531, 2017, pp.53-68

DOI:10.1007/978-3-319-48036-7_5

[学会発表] (計 8 件)

萩原良信, 井上将一, 谷口忠大: 「マルチモーダル情報に基づく階層的場所概念形成」第 31 回人工知能学会全国大会 (JSAI2017), 2017 年.

磯部匠汰, 谷口彰, 萩原良信, 谷口忠大: 「物体の共起性に基づいたマルチモーダル場所領域学習による場所の理解」第 31 回人工知能学会全国大会 (JSAI2017), 2017 年.

浅田和弥, 萩原良信, 谷口忠大, 稲邑哲也: 「没入型インタフェースを伴うクラウド仮想環境における HRI に基づく場所の学習」インタラクション 2017 年.

Yoshinobu Hagiwara, Masakazu Inoue, Tadahiro Taniguchi, “Place Concept Learning by hMLDA Based on Position and Vision Information”, 13th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems, 2016.

石伏智, 谷口彰, 萩原良信, 谷口忠大: 「マルチモーダル情報による場所領域と場所の名前の同時推定」計測自動制御学会 システム・情報部門 学術講演会 2016.

浅田 和弥, 萩原良信, 谷口 忠大, 稲邑 哲也: 「クラウド仮想環境におけるマルチモーダル情報に基づいた場所概念の形成」計測自動制御学会 システム・情報部門 学術講演会 2016.

Yoshinobu Hagiwara, “Formation of hierarchical place concept by hMLDA”, NII Shonan Meeting Seminar 092: Cognitive

development and symbol emergence in humans and robots, 2016.

福井隆士, 石伏智, 萩原良信, 谷口忠大: 「移動ロボットによる場所概念に基づく指示領域探索」第 34 回日本ロボット学会学術講演会, 2016 年.

(その他)

RoboCup@Home のチーム Web ページ

<https://emlab.jimdo.com/>

研究概要紹介 Web ページ

<https://hagiwara-robotvision.jimdo.com/research/>

6. 研究組織

(1) 研究代表者

萩原 良信 (HAGIWARA, Yoshinobu)

立命館大学・情報理工学部・助教

研究者番号: 20609416

(2) 研究協力者

谷口 忠大 (TANIGUCHI, Tadahiro)

立命館大学・情報理工学部・教授