

機関番号：12612
 研究種目：基盤研究（C）
 研究期間：2008年度～2010年度
 課題番号：20500179
 研究課題名（和文）マルチモーダルカテゴリ分類に基づくロボットの実世界理解に関する研究
 研究課題名（英文）Research on Real World Understanding by Robots Based on Multimodal Categorization
 研究代表者
 長井 隆行（NAGAI TAKAYUKI）
 電気通信大学・大学院情報理工学研究科・准教授
 研究者番号：40303010

研究成果の概要（和文）：本研究では、ロボットが本当の意味で物事を“理解”するためにはどうすればよいか、という観点からロボットの知能の問題に取り組んだ。ここでは物事に対する理解を、過去の経験をカテゴリ分類し、そのカテゴリを通じた予測であると定義した。従って、理解の基本となるカテゴリを形成するために、ロボットが知覚情報を自らカテゴリ分類する能力が必要となる。本研究では、このカテゴリ分類を教師なしで行うことのできるアルゴリズムを確立し、ロボットによる真の理解の実現可能性を示した。

研究成果の概要（英文）：In this research project, we address the problem of “robot intelligence” from a viewpoint of how robots can truly understand things. Here, we define the “true understanding” as prediction through categories (concepts), which are formed by categorizing past experiments in an unsupervised way. Therefore, the unsupervised multimodal categorization algorithm is required for robots in order to form the basis of understanding. We have developed some multimodal categorization algorithms based on statistical clustering methods and showed the possibility of the truly understanding by robots.

交付決定額

（金額単位：円）

| | 直接経費 | 間接経費 | 合計 |
|--------|-----------|-----------|-----------|
| 2008年度 | 1,500,000 | 450,000 | 1,950,000 |
| 2009年度 | 1,000,000 | 300,000 | 1,300,000 |
| 2010年度 | 1,100,000 | 330,000 | 1,430,000 |
| 年度 | | | |
| 年度 | | | |
| 総計 | 3,600,000 | 1,080,000 | 4,680,000 |

研究分野：知能ロボティクス、マルチメディア信号処理
 科研費の分科・細目：情報学、知覚情報処理・知能ロボティクス
 キーワード：知能ロボティクス、機械学習、パターン認識、人工知能

1. 研究開始当初の背景

近年、ロボットの技術的進歩は目覚ましい。しかし、ロボットの一般家庭への普及は必ず

しも進んでいるとは言えず、各家庭に一台の時代はまだ遠い未来のことに思える。これは、ロボットのメカ的な側面が注目され

る一方で、知能面での進歩が遅れていることが一つの大きな要因であると言える。ロボットの知能は、自分がおかれた環境・状況を判断し、適切な行動を取るために大変重要である。ここで行動とは、体の動きだけでなく、発話やユーザへの情報提示など、ロボットの活動全てを指す。従って、ユーザの発話に対して返答するといった対話などのコミュニケーションも含まれる。従来の知能（状況判断）は、辞書やルールなど事前に設定した論理構造に従ったマッチングによって行うことが多かった。これは、想定した世界の中であれば十分に機能するが、そうでないものに対しては全く対処できないという問題が存在する。当然のことながら、実世界の全てを事前に想定しておくことは不可能である。また、学習するアプローチをとる場合も、人間が正解の情報を正しく与えることができる状況を想定することが多い。本研究では、ロボットが本当の意味で物事を“理解”するためにはどうすればよいか、という観点からこの問題に取り組む。ここでまず重要なことは、“理解”の定義である。本研究では、物事に対する理解を、過去の経験をカテゴリ分類し、そのカテゴリを通じた予測であると定義する。例えば、コップを見た場合の理解とは、そのオブジェクトから、液体を入れたり、飲むといった行為を予測できることである。この予測は、過去に他人があるコップに水を注いで飲んだり、自分自身がそのような経験することで形成される“コップ”というオブジェクトカテゴリ（=概念）に基づいている。従って、理解の基本となるのは、オブジェクトやイベントなどの経験に基づくカテゴリであり、ロボットが得ることのできる全ての感覚情報を駆使して、自らカテゴリ分類する能力が必要となる。本研究では、このカテゴリ分類を自ら（教師なしで）行うことのでき

るアルゴリズムの確立と、ロボットへの実装による「ロボットの真の理解の実現」を目的とする。

2. 研究の目的

本研究の究極の目的は人間のように物事を理解するロボットの実現であるが、本研究課題の期間内ではその基礎となる部分として次の点を明らかにする。

(1) マルチモーダルカテゴリ分類アルゴリズム

真の理解に必要な多次元のマルチモーダル情報を用いた時空間的カテゴリ分類のアルゴリズムを開発し、その性能を評価することでこれを確立する。

(2) カテゴリを利用した予測

マルチモーダルカテゴリ分類によってカテゴリ化された経験を基に、現在経験していることからの理解、つまり未来の予測を行うアルゴリズムを確立する。

(3) 理解をベースとした学習

ロボットの、理解に基づく行動選択や学習のアルゴリズムを検討し、これを確立する。また、理解のアルゴリズムを利用することで、基本的な言語獲得を実現することが可能であると考えており、このアルゴリズムを実現する。これにより、人間の発話の理解や簡単な対話（雑談）を行うことが可能となると考えている。但し、文法などの高度なものに関しては対象としない。

(4) ロボットへの実装と実験

上述の(1)～(3)のアルゴリズムをシミュレーションベースで開発した後、これを実際のロボットに搭載する。ロボットは本来、24時間365日活動すべきものであるが、現在多くのロボットは、そのような長期間安定して活動することが困難である。また、学習は普段の活動の中で自然に行われるべき

ものであるが、学習モードを用意することが多い。本課題では、なるべく長い期間連続で活動し、活動の中で自然に、自らカテゴリ分類と学習を行うロボットを実装することを考える。この人間と共有した環境における長期間の学習により、人間に非常に近い理解を実現できることを示す。

3. 研究の方法

本研究ではまず、アルゴリズムの理論的な検討及び、実験に使用するロボットプラットフォームの整備、検討したアルゴリズムのロボットへの実装を行う。その後は、シミュレーションとロボットプラットフォームで得られる実データとを駆使した実験と評価を繰り返すことで、システム全体を改良しながら進めて行き、最終的には実環境でリアルタイム動作するロボットを実現する。

(1) 基本アルゴリズムの検討

ここで基本アルゴリズムとは、マルチモーダル情報の時空間的カテゴリゼーションアルゴリズムを指す。まず、基礎となる空間的マルチモーダルカテゴリゼーションアルゴリズムを検討する。検討中のアルゴリズムはグラフィカルモデルをベースとした確率的手法であり、視覚・聴覚・触覚情報（マルチモーダル情報）を用いて物体を人間と同じようなカテゴリに教師なしで分類することができる可能性がある。但しカテゴリには、色や形といったものから、物体のカテゴリなど様々な粒度があり、また、これらは単純な階層構造をもっているわけではない。こうしたことを考慮したカテゴリ分類アルゴリズムの開発が必要である。一方、時間的なカテゴリゼーションとしては、ニューラルネットワークを用いた手法を検討中であり、これにより、時系列データの分割とカテゴリ分類を実現できることを、音声信号を使用した実験で

確認済みである。

(2) ロボットプラットフォームの整備

実験及び最終的に提案するアルゴリズムを実装するためのロボットプラットフォームを整備する。ロボットは、上半身型のヒューマノイドロボットであり、手で物体を掴むことや、首の回転、顔の表情表出などが可能である。また、視覚・聴覚・触覚の各センサをもち、それらの信号は、並列処理されTCP/IPで通信される。現在さらに嗅覚及び味覚センサの構築を進めており、これらのモジュールを含めた全体のシステムが、完全に動作しかつ、新たなモジュールの追加や既存モジュールの保守が容易となるシステム構築を行う。

(3) 応用アルゴリズムの検討

ロボットが実際に経験することによって得るマルチモーダル情報を、カテゴリ分類によって得られたカテゴリ情報から予測するアルゴリズムを構築する。これは本課題で定義した“理解”のアルゴリズムである。さらに、この理解に基づき、行動の選択や調整を行うアルゴリズム、概念である各カテゴリと音声ラベルとの結びつきによる言語獲得アルゴリズムを検討する。アルゴリズムを検討する段階では、シミュレーションデータを用いる。

(4) 長時間駆動ロボットの構築

ロボットは、ある特定のタスクを短い時間で実験的にこなすだけでなく、日常的に活動できるように、数ヶ月間の長期に渡る駆動に耐え得るように改良する。また、全方位台車による移動機構をもたせることで、より柔軟に学習できるロボットプラットフォームを構築する。

(5) リアルタイム動作ロボットの実現

研究期間の最後では、成果の全てをロボットに搭載し、リアルタイムで全てが動作する

ロボットを実現する。また、得られた結果を評価し、研究の総括として論文執筆等を行う。

4. 研究成果

上記目的のために、初年度はまず基本となる概念の形成アルゴリズムの検討と、実装するためのロボットプラットフォームの整備を行った。

基本アルゴリズムとしては、自然言語処理で研究されている統計的クラスタリング手法である、pLSA と LDA をマルチモーダルに拡張したマルチモーダル pLSA (MpLSA) とマルチモーダル LDA (MLDA) を開発した。様々な実験を行った結果、MLDA は分類精度や予測性能において MpLSA よりも優れていることが分かり、これを基本的なアルゴリズムとして利用することとした。ただし、この段階ではカテゴリ数を自動的に推定できないという問題点があった（これについては、後に改善することができる）。また、時間的なカテゴリ分類としては、実際にロボットのセンサから得られる情報とその結果として得られる内部状態を時間方向でカテゴリ化する手法を検討し、実際に RNN を用いることでこれが可能であることを明らかにした。

ロボットプラットフォームに関しては、従来の上半身型のヒューマノイドを改良し、移動型の双腕ロボットへと発展させた。これにより、移動を含めたより複雑な行動をとることが可能となり、その後の研究の発展が見込まれた。

2年目は、昨年度に開発した概念の形成アルゴリズムを発展させ、語意の獲得・理解を実現した。これを実際のロボットに搭載し、実際にユーザとロボットの間で、理解に基づいた簡単な対話の実現が可能であることを

示した。また、このアルゴリズムを移動型ロボットプラットフォームに実装し、ロボカップ@ホームの競技で実演した。これにより、日本大会で優勝、世界大会で準優勝という成績を残すことができた。

最終年度ではまず、課題であった、オンラインでのインクリメンタルな学習と、様々な粒度での概念構造の実現を目標として取り組んだ。これらの課題を、Bag of Multimodal LDA (BoMLDA) という枠組みを提案することで解決した。これは、大量の MLDA を様々な重み付けられてマルチモーダルデータにより生成し、その中から有用なモデルを選択するというものであり、選択の基準としてもユーザとのインタラクションや予測性能の推定など様々なものが利用できる。また、アルゴリズムをオンライン化することに成功し、ロボットは、様々な粒度の複雑な概念と語意を、ユーザとのインタラクションを通して、オンラインかつインクリメンタルに学習することが可能となった。さらに、ノンパラメトリックベイジアルゴリズムを応用したマルチモーダル HDP (MHDP) により、カテゴリ数を自動的に推定可能なカテゴリ分類手法も開発した。

これらのアルゴリズムをロボットに実装し、前年度同様、ロボカップ@ホームにおいて評価した。その結果、日本大会、世界大会共に優勝という大きな成果を上げることができた。

こうした確率モデルをベースとしたカテゴリ分類アルゴリズムを実際のロボットに実装し評価することで、本研究で定義した“理解”の実現可能性を示すことができたと考えている。

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 6 件)

- ① T. Nakamura, K. Sugiura, T. Nagai, N. Iwahashi, T. Toda, H. Okada, T. Omori, Learning Novel Objects for Extended Mobile Manipulation, Journal of Intelligent and Robotic Systems, accepted
- ② T. Nakamura, T. Araki, T. Nagai, N. Iwahashi, Grounding Word Meanings in LDA-Based Multimodal Concepts, Advanced Robotics, accepted
- ③ 中村友昭, 長井隆行, 機能と視覚的情報の関係性に基づく物体の概念モデル, 情報処理学会論文誌, Vol. 51, 1413-1427, 2010, 査読あり
- ④ 中村慎也, 岩橋直人, 長井隆行, 実世界における人とロボットの共有信念の推定に基づいた適応的な発話生成, 知能と情報, Vol. 16, No. 5, 663-682, 2009, 査読あり
- ⑤ 中村慎也, 長井隆行, 岩橋直人, 佐藤健, 麻生英樹, 変分ベイズ法を用いたモデル構造探索に基づく抽象的単語の学習, 電子情報通信学会論文誌D, vol. J92-D, No. 4, 467-479, 2009, 査読あり
- ⑥ 中村友昭, 長井隆行, 岩橋直人, ロボットによる物体のマルチモーダルカテゴリゼーション, 電子情報通信学会論文誌D, Vol. J91-D, No. 10, 2507-2518, 2008, 査読あり

[学会発表] (計 13 件)

- ① 中村友昭, 長井隆行, 岩橋直人, Gibbs Samplingによる物体のマルチモーダルカテゴリゼーション, 日本ロボット学会学術講演, 名古屋工業大学, 2010年9月23日

- ② 長井隆行, マルチモーダルカテゴリゼーションに基づくロボットの概念形成と実世界理解, 電子情報通信学会ソサエティ大会 (招待講演), 大阪府立大学, 2010年9月16日
- ③ 中村友昭, 長井隆行, 岩橋直人, 複数のマルチモーダルLDAを用いた抽象的概念の形成, 人工知能学会全国大会2010, 長崎, 2010年6月10日
- ④ T. Nakamura, T. Nagai, Object Concept Modeling Based on the Relationship among Appearance, Usage and Functions, IEEE/RSJ International Conference on Intelligent Robots and Systems, Taiwan, 21Oct. 2010
- ⑤ M. Attamimi, A. Mizutani, T. Nakamura, T. Nagai, K. Funakoshi, M. Nakano, Real-Time 3D Visual Sensor for Robust Object Recognition, IEEE/RSJ International Conference on Intelligent Robots and Systems, Taiwan, 20Oct. 2010
- ⑥ M. Attamimi, A. Mizutani, T. Nakamura, K. Sugiura, T. Nagai, N. Iwahashi, H. Okada, T. Omori, Learning Novel Objects Using Out-of-Vocabulary Word Segmentation and Object Extraction for Home Assistant Robots, IEEE International Conference on Robotics and Automation 2010, Anchorage, 6May2010
- ⑦ 中村友昭, 西田匡, 長井隆行, 把持動作による物体カテゴリの形成と認識, 情報処理学会全国大会, 東京大学, 2010年3月11日
- ⑧ 中村友昭, 長井隆行, 岩橋直人, マルチモーダル情報に基づく物体概念の形成, 電子情報通信学会研究会, 京都大学, 2010年1月21日

- ⑨ T.Nakamura, T.Nagai, N.Iwahashi, Grounding of Word Meanings in Multimodal Concepts Using LDA, IEEE/RSJ International Conference on Intelligent Robots and Systems 2009, St.Louis, 13 Oct. 2009
- ⑩ 中村友昭, 長井隆行, 岩橋直人, マルチモーダルな物体概念に基づく語彙の理解, 日本ロボット学会学術講演会, 横浜国立大学, 2009年9月16日
- ⑪ 中村友昭, 長井隆行, 岩橋直人, LDAを用いたマルチモーダル物体概念に基づく語意獲得, 電子情報通信学会総合大会, 愛媛大学, 2009年3月19日
- ⑫ 中村友昭, 長井隆行, 岩橋直人, マルチモーダルLDAを用いたロボットによる物体のカテゴリゼーション, 日本ロボット学会学術講演会, 神戸大学, 2008年9月10日
- ⑬ 中村友昭, 新地康人, 長井隆行, グラフィカルモデルを用いた物体概念モデル, FIT2008第7回情報科学技術フォーラム, 慶應義塾大学, 2008年9月3日

[図書] (計1件)

- ① T.Nakamura, T.Nagai, Forming Object Concept Using Bayesian Network, Bayesian Network, Ahmed Rebai (Ed.), ISBN: 978-953-307-124-4, Sciyo, 2010.08

[産業財産権]

○出願状況 (計1件)

名称: 触覚 (手触り) 情報の取得装置と記録方法

発明者: 長井隆行, 中村友昭, 石井香澄

権利者: 長井隆行, 中村友昭, 石井香澄

種類: 特願

番号: 2008-210075

出願年月日: 平成 20 年 8 月 18 日

国内外の別: 国内

6. 研究組織

(1) 研究代表者

長井 隆行 (NAGAI TAKAYUKI)

電気通信大学・大学院情報理工学研究科・准教授

研究者番号: 40303010

(2) 研究分担者

なし

(3) 連携研究者

なし