

機関番号：82636

研究種目：基盤研究（C）

研究期間：2008～2010

課題番号：20500186

研究課題名（和文）生活支援ロボットの対話と行動のユーザー適応化技術の研究

研究課題名（英文）A Study on User Adaptation of Dialog and Behavior for Home Assistant Robots

研究代表者

岩橋 直人（IWAHASHI NAOTO）

独立行政法人情報通信研究機構・知識創成コミュニケーション研究センター音声コミュニケーショングループ・専攻研究員

研究者番号：90394999

研究成果の概要（和文）：生活支援ロボットとの安全・安心なコミュニケーションを実現するために、ロボットの対話と行動の機能をユーザの生活空間と行動へ適応させる技術として、（1）物探しを指示する発話の学習と理解の手法、（2）行動指示発話の学習と理解の手法、（3）ユーザとの共有信念を効率的に形成するための発話生成手法、を開発した。さらに、これらの技術を実装した自律型ロボットで、学術的国際競技会である RoboCup@Home に出場し、優勝を果たした。

研究成果の概要（英文）：In order to realize the safe and secure communication with home assistant robots, we developed the following three methods which enable the robots to adapt their dialog and behavior to user's life space and behavior: (1) learning and understanding of mobile manipulation direction utterances, (2) learning and understanding of motion direction utterances, and (3) utterance generation for forming mutual beliefs with users. Moreover, using the autonomous robot that was implemented these methods to, we won first prize at RoboCup@Home which is an international competition for home assistant robots.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2008年度	1,300,000	390,000	1,690,000
2009年度	1,300,000	390,000	1,690,000
2010年度	900,000	270,000	1,170,000
総計	3,500,000	1,050,000	4,550,000

研究分野：人工知能

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：知能ロボット

1. 研究開始当初の背景

介護や福祉など日常生活を支援するロボットは、実用化を視野に入れて世界中で活発に研究が進められている。例えば、生活支援ロボットの開発を促進するために、日常的

活空間での自律的行動や人間-ロボットインタラクションの技術向上を目的にした学術的競技会である「RoboCup@Home」の第1回世界大会が2007年7月にアメリカ・アトランタで開催され、世界中から多くの大学の

研究者が参加した。日常生活支援ロボットでは、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、日常生活支援ロボット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、ユーザ毎に異なる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、岩橋（研究代表者）は、予め言語知識を与えるのではなく、対話と行動のコミュニケーション能力を、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ユーザとのコミュニケーションを通して学習する計算機構（以後、L-Core）を開発した。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することが可能である。本技術を実装したロボットが、ユーザとのコミュニケーションを通して、数十個の単語と単純な文法などを学習し、ユーザの発話に従い、状況に応じて適切に机の上に置いた縫いぐるみを操作したり、質問に音声で答えたりできるようになることを実証した。

2. 研究の目的

本研究課題では、上記の計算機構 L-Core を、机の上の空間でのコミュニケーションから、日常的生活空間であるリビングルームでのコミュニケーションで利用できるように拡張する。ロボットが、ユーザの生活空間だけでなく生活習慣をも考慮して、「いつものあれ持ってきて」などといった、自然であいまいな発話による指示を適切に理解して、その状況でユーザが必要としているカバンなどを持って来ることができるようにする。

3. 研究の方法

サブゴール 1, 2 を設定し、3 年の研究期間のうち、1, 2 年目にそれらを順番に達成してゆき、最終年に本研究の目的を達成した。

サブゴール 1, 2 は以下のとおりである。

サブゴール 1. 物探しを指示する発話の学習と理解: ロボットに、コップやカバンなどの物体を手を持って見せてその名前を音声で教える。名前はユーザが自由に決めることができる。人間が、教えた物の一つを部屋のどこかに置き、「〇〇を探して」という指示をする。ロボットは部屋の中を探して、それを見つけて人間に知らせる。

サブゴール 2. 行動指示発話の学習と理解: 実際に人間がロボットにやって見せることで、物体操作を教える。同時に、その操作を指示する単語を音声で教える。人間が教えた単語を使ってロボットに操作指示を出し、ロボットがそれに従い行動する。具体的には、“～を～に乗せる”, “～を～に入れる”, などの物体操作を学習させる。各動作において、ユーザの生活空間や状況に応じて、適切な軌道、速度、把持の仕方、置き方などの制御法が学習される。操作指示は、例えば、「大きなコップをテーブルの上に乗せて」などである。

4. 研究成果

(1) 日常生活支援ロボットの音声対話技術に着目すると、既存のトップダウン手法においては言語知識があらかじめ与えられていることが多い。しかし、全ての言語知識を用意することは不可能であるため、トップダウン手法には未登録語を発声できないという制約がある。例えば、案内ロボットがユーザの顔画像を学習できたとしても、人名が登録されていなければ名前を呼ぶことはできない。一方、画像と音声からボトムアップに語彙を獲得するロボットの研究も行なわれているが、実用性に問題がある。

これに対し我々は、言語知識を利用しつつ、未登録語の発声が可能なハイブリッド手法を開発した。具体的なタスクとしては、新規物体の学習を扱う。我々の提案する手法では、未登録語の登録をテンプレート文で行ない、通常の対話はルールベースで行なう。ユーザがロボットに物体を見せながら「この名前は X」と発話することにより物体を学習させ、ロボットに再度物体を見せた場合に「これは X です」と発話させるようなタスクである (図 1)。



図 1 新規物体学習タスクにおけるインタラクション

いま、新規物体に関する画像とテンプレート文で発話された音声を与えられたとする。このとき、新規物体の学習のために解くべき問題は以下の 4 つに分類できる。すなわち学習時には、1) 雑音下での頑健な音声認識、2) 画像から学習すべき物体の抽出、が必要であり、認識時には、3) 照明条件の変化に対して頑健な物体認識、4) 未登録語の発声、が必要である。これに対し我々は、動きアテンションに基づく物体抽出 (図 2)、scale-invariant feature transform (SIFT) 情報によるマッチング、ノイズの逐次推定と雑音抑圧、音声からの未登録語の切り出しと声質変換 (図 3)、を組み合わせることでこれらの問題を解決した。

音声・画像処理を統合したシステムをロボットに実装し、リビングルーム環境において実験を行なった。実験の目的は、「ある新規物体に対し、学習フェーズでは未登録語の切り出しおよび物体抽出が成功し、かつ認識フェーズでは画像の認識が成功する確率 (統合精度)」を評価することである。学習フェーズでは、ユーザが物体を見せながら未登録語 X を発声した後、ロボットが X を返答すれば成功とする。認識フェーズで、ユーザが物体を見せながら音声で名前を質問した後、ロボットが X を返答すれば認識フェーズ成功とする。登録物体全てに対し学習フェーズを行なった後、認識フェーズを行なう。本実験では、ロボットに搭載したモニタより区間検出と学習の経過情報をユーザにリアルタイムでフィードバックしている。これは、発話中に区間検出誤りに関する情報をユーザに与えることで、区間検出精度を改善できるためである。学習フェーズでユーザの言い直しがなく成功した場合と、未登録語切り出しの失敗のためにユーザが言い直した場合について統合精度を調べた。実験の結果、統合精度は 88%であった。学習フェーズにおいて言い直しを許した場合、統合精度は 94%であった。

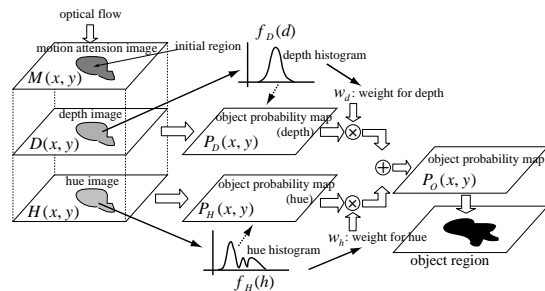


図 2 動きアテンションに基づく物体抽出

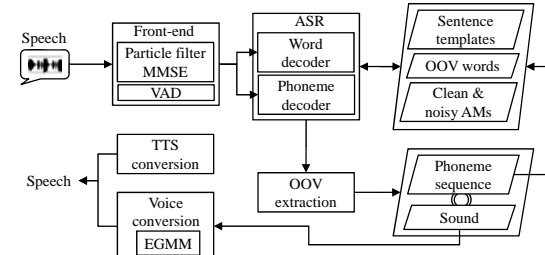


図 3 音声処理の概要

(2) 行動指示発話の学習と理解の手法を開発した。本手法によれば、実際に人間がロボットにやって見せることで、模擬的な家の中の生活環境で物体操作を教えることができる。同時に、それらの行動を指示する単語を音声で教え、教えた単語を使ってロボットに動作指示を出し、ロボットがそれに従い行動することができる。

ユーザが数回、物体操作動作を見せるだけで動作の学習を行う手法を、車輪型双腕付き移動ロボット (図 4) に実装した。上記 (1) で述べた、物探しを指示する発話の学習と理解の手法を併用し、ユーザが音声で指示をすることで、部屋の中にある物体を見つけて、指示通りの操作をすることを可能とした。

さらに、ユーザの物体指示発話があいまいであった場合に、あいまい性解消のための発話を効用理論に基づいて生成する手法を開発し、ロボットが誤ってユーザの意図しない動作を行なうことを未然に防ぐことができる安全・安心なインタラクションを実現した。

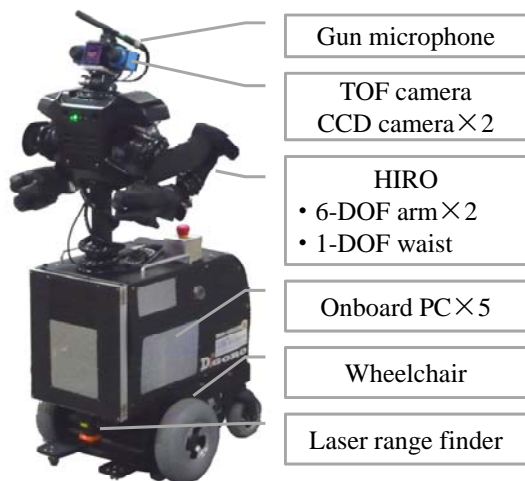


図 4 開発した手法を実装したロボット

(3) 従来の対話技術には安全性上の観点から大きな問題がある。それは、ユーザの発話の意味が適切に理解されずに、ロボットが予期しない動作を行ってしまう危険性があることである。この危険性を減少させることを目的とした手法の開発を行った。具体的なタスクとして、物体操作対話タスクを扱った。物体操作対話タスクにおいて、ユーザの発話の意味が適切に理解されるためには、言語による動作参照、および言語によるオブジェクト参照における曖昧性を解消する必要がある。

能動学習の枠組みにより発話理解確信度を学習する手法を開発した。本手法では、確信度関数の学習に必要な発話をロボットに選択させ、その発話を受けてユーザが物体を操作する。発話の選択には、能動学習の一種である Expected Log Loss Reduction (ELLR) を用いた。提案手法の独自性は、i) 対話システムへの能動学習の導入、ii) 確信度の学習に Bayesian Logistic Regression (BLR) を用いること、の 2 点である。これまで一般の対話システムでは、ユーザの発話を受けてシステムが応答を生成し、ユーザ発話を理解するモデル(ユーザモデル)を学習させることは可能であったものの、システムの発話によりユーザモデルを学習させることは困難であった。一方、本手法では、ロボットとユーザが実世界にグラウンドした対話を行うため、ロボットが「ユーザが用いる語彙」を用いてユーザに動作応答させることができる。このような特徴から、対話戦略として能動学習を用いることが初めて可能になった。

発話理解確信度を表す関数の能動学習に関する定性的結果を図 5 に示す。図における

(a) から (c) の曲線は、それぞれ訓練サンプル数 10、20、30 のときの BLR による回帰結果を表す。図より、サンプル数 30 において収束していることがわかる。図 6 は、エピソード数(ロボットの発話数)に対する、選択されたマージン(発話の曖昧性の低さを表す数値)の変化を示す。図において、点線はユーザの動作が失敗したエピソードを示す。図より、ユーザの動作失敗の次のエピソードでは、より大きいマージンを持つ(曖昧性の低い)発話を選択されていることがわかる。

本手法では、能動学習の枠組みにより確信度の学習に有効な発話を選択される。音声合成されたこれらの発話と、発話に対するユーザの動作応答から、「ユーザの発話に対する確信度関数」が学習される。このようにして学習された確信度関数を用いることにより、実行フェーズでの動作失敗率についても減少させることが可能である。

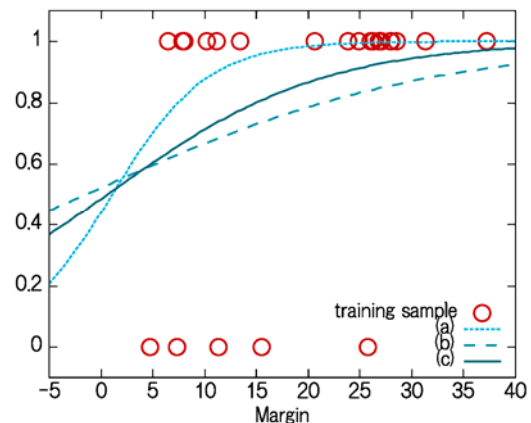


図 5 発話理解確信度を表す関数の学習結果。○はサンプルを表す。

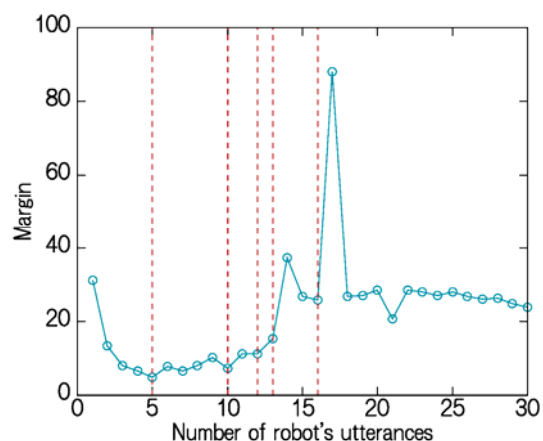


図 6 選択された発話のマージンの変化

(4) 上記 (1)、(2)、(3) などの手法を自律型双腕移動ロボットに実装し、ロボカップ@ホームリーグに出場した。2008 年は、日本

大会、世界大会ともに優勝。2009年は、日本大会優勝、世界大会準優勝。2010年は、日本大会、世界大会ともに優勝。20ほどのチームが参加した世界大会でも、圧倒的な強さを発揮した。

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 11 件)

- ① Tomoaki Nakamura, Takaya Araki, Takayuki Nagai, Naoto Iwahashi, "Grounding of Word Meanings in LDA-Based Multimodal Concepts," *Advanced Robotics*, 査読有, 掲載決定.
- ② Komei Sugiura, Naoto Iwahashi, Hideki Kashioka, and Satoshi Nakamura. "Learning, Generation, and Recognition of Motions by Reference-Point-Dependent Probabilistic Models," *Advanced Robotics*, 査読有, Vol. 25, No. 6-7, 2011, pp. 825-848.
- ③ 杉浦孔明, 岩橋直人, 柏岡秀紀, 中村哲. "言語獲得ロボットによる発話理解確率の推定に基づく物体操作対話確認発話," *日本ロボット学会誌*, 査読有, Vol. 28, No. 8, 2010, pp. 978-988.
- ④ 岡田浩之, 大森隆司, "ロボカップ@ホーム一人とロボットの共存を目指して," *人工知能学会論文誌*, 査読有, Vol. 5, No. 2, 2010, pp. 229-236.
- ⑤ 谷口忠大, 岩橋直人 "複数予測モデル遷移の N-gram 統計に基づく非分節運動系列からの模倣学習手法", *知能と情報 (日本知能情報ファジィ学会誌)*, 査読有, Vol. 21, No. 6, 2009, pp. 1143-1154.
- ⑥ 中村慎也, 岩橋直人, 長井隆行, 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*, 査読有, Vol. J92-D, No. 4, 2009, pp. 467-490.
- ⑦ 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*, 査読有, Vol. J91-D, No. 10, 2008, pp. 2507-2518.
- ⑧ Taniguchi, T., Iwahashi, N., Sugiura, K., Sawaragi, T. "Constructive Approach to Role-Reversal Imitation Through Unsegmented Interactions," *Journal of Robotics and Mechatronics*,

査読有, Vol. 20, No. 4, 2008, pp. 567-577.

[学会発表] (計 54 件)

- ① Iwahashi, N., "A Developmental Approach to Situated Dialog with Robots," *Workshop on developmentl speech recognition*, 2011年2月17日, Center of Excellence Cognitive Interation Technology, Bielefeld, Germany.
- ② Komei Sugiura, Naoto Iwahashi, Hideki Kashioka, Satoshi Nakamura. (2010) "Active Learning of Confidence Measure Function in Robot Language Acquisition Framework," *Int. Conf. Intelligent Robots and Systems*, 2010年10月18日, Taipei International Convention Center, Taipei, Taiwan ROC
- ③ Muhammad Attamimi, Akira Mizutani, Tomoaki Nakamura, Komei Sugiura, Takayuki Nagai, Naoto Iwahashi, Hiroyuki Okada, and Takashi Omori. (3 May 2010). "Learning Novel Objects Using Out-of-Vocabulary Word Segmentation and Object Extraction for Home Assistant Robots", *Proc of IEEE Int. Conf. Robotics and Automation*, 2010年5月3日, Anchorage Convention District, Alaska, USA.
- ④ 板谷純希, 中村友昭, 長井隆行, "自然言語によるロボットの行動プログラミング," *情報処理学会全国大会*, 2010年3月11日, 東京大学, 東京都
- ⑤ 杉浦孔明, 水谷了, 中村友昭, 長井隆行, 岩橋直人, 岡田浩之, 大森隆司, "音声からの未登録語切り出しと画像からの物体抽出の統合による新規物体の学習", 第26回日本ロボット学会学術講演会, 2008年9月9日, 神戸大学, 兵庫県
- ⑥ 岡田浩之, 大森隆司, 岩橋直人, 長井隆行, 杉浦孔明, "ロボカップ@ホームにおける音声対話技術", *人工知能学会全国大会*, 2008年6月13日, ときわ市民ホール, 北海道

[図書] (計 3 件)

- ① Sugiura, K., Iwahashi, N., Kashioka, H., and Nakamura, S. "Statistic Imitation Learning in Sequential Object Manipulation Tasks" in *Ernest Hall (ed.)*, *Advances in Robot*

Manipulators, pp. 589-606, INTECH, 2010.

- ② Taguchi, R., Iwahashi, N., and Nitta, T. "Learning Communicative Meanings of Utterances by Robots," In Hattori, H., Kawamura, T., Ide, T., Yokoo, M., Murakami, Y. (eds.), New Frontiers in Artificial Intelligence, LNCS / LNAI 5447, pp. 62-72, Springer, 2009.

[産業財産権]

○出願状況 (計 1 件)

名称：情報処理装置、情報手法、プログラム及び対話システム

発明者：杉浦孔明, 岩橋直人

権利者：(独)情報通信研究機構

種類：特願

番号：1009-204599

出願年月日：2009年9月4日

国内外の別：国内

○取得状況 (計 0 件)

6. 研究組織

(1) 研究代表者

岩橋 直人 (IWAHASHI NAOTO)

独立行政法人 情報通信研究機構・知識創成コミュニケーション研究センター音声コミュニケーショングループ・専攻研究員
研究者番号：90394999

(2) 研究分担者

長井 隆行 (NAGAI TAKAYUKI)

電気通信大学・大学院電気通信研究科・准教授

研究者番号：40303010

岡田 浩之 (OKADA HIROYUKI)

玉川大学・工学部・教授

研究者番号：10349326

杉浦 孔明 (SUGIURA KOMEI)

独立行政法人 情報通信研究機構・知識創成コミュニケーション研究センター音声コミュニケーショングループ・専攻研究員
研究者番号：60470473

(3) 連携研究者

()

研究者番号：