

**科学研究費助成事業 研究成果報告書**

平成 28 年 6 月 20 日現在

機関番号：13903

研究種目：若手研究(B)

研究期間：2013～2015

課題番号：25730111

研究課題名(和文) サービスロボットによる実インタラクションデータからの語彙学習

研究課題名(英文) Lexical learning from interaction data for service robots

研究代表者

田口 亮 (Taguchi, Ryo)

名古屋工業大学・工学(系)研究科(研究院)・助教

研究者番号：70508415

交付決定額(研究期間全体)：(直接経費) 3,200,000円

研究成果の概要(和文)：ロボットが人の命令を理解し動作するためには、認識可能な単語や文法、および、それらに対応する行動を事前に設計する必要がある。しかし、事前知識に無い単語が発話された場合、ロボットは動作できない。そこで本研究ではインタラクションを通してユーザの言い回しを学習し、既存の知識を拡張させていく語彙学習手法を開発した。また、移動指示の命令では相対的位置概念が用いられる。このような概念を学習するためには、基準となる物体(参照対象)を推定する必要がある。本研究では、ロボットが自律的に獲得した地図情報から、参照対象候補を抽出し、参照対象の推定と位置概念の学習を同時に行う手法を提案した。

研究成果の概要(英文)：In order to understand speech of a user and act adequately, service robots need to have the pre-designed linguistic knowledge, such as lexicon, grammar and the relationship between the meaning of speech and actions of robots. However, the robot cannot adequately act when a user uttered the term that is not included in the pre-designed knowledge. To solve this problem, in this research, we developed the method to extend the linguistic knowledge by statistically analyzing phoneme sequences of the new wordings. On the other hand, utterances for navigation often includes relative spatial concepts. However, reference objects which are required to understanding relative spatial concepts may not be included in the utterances. In this research, we proposed a method for learning relative spatial concepts while estimating the reference objects. In this method, a robot extracts candidates of reference object from an occupancy grid map.

研究分野：情報工学

キーワード：音声情報処理 言語獲得 シンボルグラウンディング ロボット 発話理解

## 1. 研究開始当初の背景

ロボットが人と対話するためには、言葉と実世界の事物・事象の対応関係をロボットが理解できなければならない。家庭やオフィスなどでは、未知の人や物、場所等に対応する必要があるため、それらを表す単語知識、すなわち語彙をユーザとのインタラクションを通して学習できることが望まれる。

ロボットによる語彙学習に関する先行研究では、人がロボットに物や動作を見せながら対応する単語を発話することで、「箱」や「青い」といった物を表す単語や、「乗せて」や「近づけて」といった動作を表す単語を学習させた。また、自由な言い回しの発話から語彙を学習する手法も開発されている。

これらの先行研究で開発された手法は単語の事前知識を用いないため、任意の単語を学習させることができるが、1つの単語の音韻系列を獲得させるためには10回程度の音声教示を必要とした。しかし、実際にサービスロボットが語彙を学習する場面を想定すると、未知の言い回しが発話された場合にのみ上記学習手法を適用し、既知の言い回しで教示された場合や、未知の言い回しを学習した後には、即時に単語の音韻系列を獲得できることが望ましい。また、従来は発話から単語を学習する学習モードと、学習結果を用いて命令発話を認識し動作する動作モードを明示的に切り替えていたが、実環境では、ユーザの発話が教示発話なのか、命令発話なのか判別する必要があり、命令発話を誤って教示データに加え学習してしまう場合が考えられる。そのため学習モードと動作モードの切り替えを必要としない学習方法が望まれる。

## 2. 研究の目的

本研究では、学習した知識や設計時に与えられた単語知識を利用した語彙学習手法と、単語知識を必要としない学習手法を併用することで、より効率的で実用的な語彙学習手法を開発することを目的とする。また、学習に用いる発話は、従来のような「これは〇〇です」というような単語教示用の発話ではなく、タスク達成に必要な命令発話を用いる。これにより学習モードと動作モードの切り替えを必要としない学習が実現する。

## 3. 研究の方法

本研究では先行研究で開発された手法を基に、単語事前に与えられる単語知識を利用した手法の開発し、従来法との統合を行う。学習用の命令発話を収集するために、計算機シミュレーションだけでなく、人の発話により実際に動作するロボットを開発する。また、実際のタスクの中での学習実験を実現するために必要な、物体認識アルゴリズム、相対的位置概念の学習アルゴリズムを開発する。

## 4. 研究成果

## (1) 語彙学習手法

命令理解のための事前知識を用いた音声認識と、ユーザの自由な言い回しから単語学習を行う手法を併用することで、未知の単語を含む命令発話から、その単語の音韻系列と対応する意味を推定する手法を開発した。提案手法では、事前知識に無い発話がなされた場合には、その発話の音素系列を保持し、ロボットは「別の言い回しで言い直して下さい」と発話し、再発話を要求する。その後、登録済みの文法で入力があると、対応する意味IDを先の未知発話の音素系列と共に記録し、学習用のデータベースを作成する。例えば、既知の文法として「右に回転して」が登録されており、「右に」が「Right」、「回転して」が「Rotate」という意味IDに対応付けられているとする。この時、「みぎのほうむいて(右の方向いて)」とユーザが発話しても、登録文法と異なるためロボットは正しく認識することができない。再発話の要求後、ユーザが正しく「右に回転して」と発話すると、ロボットは先の「みぎのほうむいて」に対応する意味ID列は「Right Rotate」であると判断できる。しかし、発話のどこまでが「Right」で、どこからが「Rotate」に対応するかわからない。そこで、まず、全学習データの音素系列を教師なし形態素解析手法により形態素に分解する。形態素解析の結果は、「みぎ／の／ほう／むいて」というように、意味IDの単位よりも短く分節される傾向がある。そこで、全学習データを用いて各形態素を条件とした意味IDの条件付き確率を求め、各発話で意味ID列の尤度が最大となるように形態素を連結し単語を獲得する。

提案手法の評価は2種類のロボットを用いて実施した。実験1では、移動ロボットを用いて実験を行った。実験の様子を図1に示す。音声認識器にはJuliusを利用した。初期に与える文法と、対応する意味ID列を表1に、学習させる新規語を表2に示す。実験では1名の話者がロボットとインタラクションを行い、表2の新規語を学習させた。インタラクション中の発話は34回、その内、新規語を用いた発話は13回あった。学習の結果、得られた語彙を表3に示す。

音素認識誤りにより、既存の言い回しの“まえ”と“ひだり”が、“わえ”、“ひでる”と異なる新規語として獲得された。新規語に対応する単語は計16個獲得された。本来ロボットに学習させるべき単語は10個であり、6個多くの単語を学習している。これは、“みて”と“みって”、“にかえてんしてう”と“にかいってして”のような音素認識のゆらぎによるものである。提案手法では、類似の単語をマージする機能が無いため、このようなゆらぎを吸収することはできない。しかし、想定した新規語は全て1回～3回程度の発話から獲得できて



図 1: 実験 1 の様子

表 1: 実験 1 の初期知識

単語	意味 ID	意味 ID 列
まえ	FORWARD	FORWARD MOVE
うしろ	BACK	BACK MOVE
みぎ	RIGHT	RIGHT ROTATE
ひだり	LEFT	LEFT ROTATE
にむかって	MOVE	
のほおむいて	ROTATE	

表 2: 実験 1 で学習させる新規語

FORWARD	BACK	RIGHT	LEFT
ぜんぼう	こうぼう	みぎがわ	ひだりがわ
MOVE		ROTATE	
にすすんで		にかいてんして	
にいて		みて	
のほおにいて		のほおみて	

表 3: 実験 1 で獲得された語彙

FORWARD	BACK	RIGHT	LEFT
まえ	うしろ	みぎ	ひだり
わえ	うしろごほ	みぎが おわ	ひでる
ばえのほお	こほお		ひだりがわ
ぜんほお			
ぜんほおい			
MOVE		ROTATE	
むかって		むいて	
にいて		みて	
にすんであ		にかえてんしてう	
にいて		のほおみて	
にすねう		みて	
		にかいてんして	

おり、従来よりも効率的な学習が実現できたと言える。

実験 2 では、キズ検査用ロボットに提案手法を実装し、評価実験を実施した。キズ検査ロボットの外観を図 2 に示す。研究で用いるロボットは、車両側面のキズを検査することを目的に開発されたロボットであり、検査のための LED 直線照明とカメラ、動作制御用のタッチパネル、PC2 台を搭載している。検査



図 2: キズ検査ロボットの外観

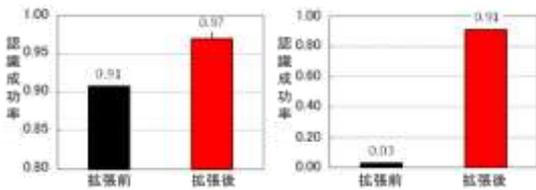
表 4: 初期知識の文法と意味 ID 列

登録文法	意味 ID 列
元画像 見せて	IMG SHOW
元画像 保存	IMG SAVE
二値画像 見せて	BINARY SHOW
二値画像 保存	BINARY SAVE
局所分散 見せて	LOCAL SHOW
局所分散 保存	LOCAL SAVE
検出結果 見せて	RESULT SHOW
検出結果 消して	RESULT OFF
反射領域 見せて	RANGE SHOW
反射領域 消して	RANGE OFF
閾値 上げて	TH UP
閾値 下げて	TH DOWN
止まって	STOP

表 5: 新規に使用する言い換え表現

意味 ID	登録語	新規語
SHOW	見せて	ひょうじ
		きりかえて
SAVE	保存	さつえい
		とって
OFF	消して	さくじょ
		なくして
UP	上げて	あつぷ
		たかく
DOWN	下げて	だうん
		ひくく
STOP	止まって	すとつぷ
		とまれ

の際には、対象に照明光を照射し、その反射光をカメラで撮影する。キズが生じている部分は周辺と法線方向が異なるため、直接反射が得られず黒く浮かび上がる。そこで、画像の局所領域で分散を求めることでキズを強調し検出する。本ロボットの運用場面では、検査対象までの移動や、検査開始姿勢の調整、カメラパラメータの調整を、検査員がタッチパネルまたは音声を用いて行う。実験では、表 4 に示す文法を登録しておき、表 5 に示す動詞部分の言い換えを学習させた。13 種の文法に対し、既知の単語を用いた発話を 1 つ、未知の単語を用いた発話が 2 つとなり、発話の種類は 39 種類となる。これを各 10 回ずつ発話し、計 390 個の音声データを収録した。収録データの内、半分を学習データ、残りを



(a) 既知発話 (b) 未知発話

図 3: 実験 2 の結果

評価データとする 2-fold cross validation により評価した. 実験結果を図 3 に示す. ユーザが意図した意味 ID 列を出力できた場合を認識成功とみなした場合の正解率が示されている. 図 3(b) から確認できるように, 単語辞書を拡張した結果, 未知発話の認識成功率が上昇している. このことから提案手法の有効性が確認できる. また, 図 3(a) より, 既知発話についても提案手法の認識成功率が高いことが確認できた. 既知の発話でも誤認識する場面があるが, 提案手法によりユーザ毎の発音のゆらぎを学習することにより誤認識が減少したためである.

本研究の目的はより効率的で実用的な語彙学習手法の開発である. 提案手法は, タスクに使用された命令発話を用いることで学習モードと動作モードを切り替えることなく学習が行えるため, 従来よりも実用性の高い語彙学習手法であると言える. また, 実験 1 で示したように, 新規語を含む言い回しでも, 1 回~3 回という少ない発話から学習が可能であり, 従来よりも効率性の高い語彙学習手法が実現できたとと言える.

## (2) 物体認識アルゴリズムの開発

ロボットに物体の名前を教えるという場面では, ロボットの目の前で人が物体を持ち, 任意に回転させながら物体の名称を教えることが想定される. その場合, 物体の姿勢は未知となり, さらに物体の動かし方に依存して得られる視覚特徴が不均一となるため識別性能が低下する. そこで本研究では, 自己組織化マップ (SOM) を用いることで, 物体の姿勢変化にロバストな物体概念学習手法を開発した. 提案手法では, 画像の局所特徴量として SIFT, 特徴量の表現に Bag of Features (BoF), 特徴群のクラスタリングに SOM, 識別器に SVM を用いている. 提案手法で学習する際の処理の流れを以下に示す.

### Step1. 特徴抽出

- ① 画像から SIFT 特徴を抽出
- ② 画像中の SIFT 特徴をヒストグラムで表現 (BoF)

### Step2. 特徴学習

- ① SOM によりヒストグラムをクラスタリング
- ② SOM のユニットごとに SVM を学習

実験の様子を図 4 に示す. また 2×2 ユニ



図 4: 物体名前を教える様子



図 5: SOM によるクラスタリングの結果



図 6: 物体認識率

ットの SOM によるクラスタリングの結果を図 5 に示す. このように提案手法では, 物体の角度が未知の場合でも, 見えの違いにより画像をクラスタリングすることができる. 20 種類の物体を用いて評価実験を実施した. 1 物体当たり 300 枚の画像を学習データ, 背景の異なる 100 枚の画像を評価データとし, 交差確認法で認識率を評価した. その結果を図 6 に示す. 図から, SOM や SVM 単体で用いる場合と比べ提案手法 (SOM+SVM) では認識率が向上していることが確認できる.

## (3) 相対的位置概念学習アルゴリズム

ナビゲーションのための発話では相対的な位置概念が用いられる. 位置概念は, 基準となる物体 (以下, 参照点と呼ぶ) との

比較により表されるため、概念を学習するためには、複数の候補から真の参照点を推定しながら学習する必要がある。

我々はこれまでにEMアルゴリズムを用いて参照点の推定と概念の学習を同時に行う手法を提案していた。しかし、参照点の候補が多数ある場合に、参照点の推定を誤るといった問題や、概念の解釈が特定の座標系に固定されており、学習できない概念が存在した。本研究では事前知識により参照点の推定誤りを軽減する手法、および複数の座標系候補から適切な座標系を推定する手法を開発した。また、従来手法では参照点の候補も事前に与えられることを仮定していた。しかし、実環境でロボットが位置概念の学習を行うためには、環境との相互作用を通して自ら作成した地図情報から参照点候補を抽出する必要がある。そこで本研究では、レーザレンジファインダを搭載した移動ロボット（図7参照）を用いて地図画像を生成し、得られた画像から参照点の候補を抽出する手法を開発した。

図8に実験環境の写真、図9にFastSLAMを用いてロボットが獲得した地図画像と抽出された参照点候補（赤色の点）を示す。現在の手法では、障害物で囲まれた空間の重心を参照点候補として抽出するため、物体が密集している場合にはそれぞれを分離して参照点候補として抽出することはできない。また、壁などのように、連続した物体の一部分を参照点候補とするはできない。この問題の解決は今後の課題とする。

得られた参照点候補を用いて相対的位置概念（手前、奥、右、左）を学習させる実験を行った。概念を表す特徴量として方向を用いた。方向は図10に示すように、話者（人間）と参照点とロボットの位置関係から算出する。部屋領域内でランダムにロボットと話者の位置を決定し、話者は任意の参照点を選択し、相対的な位置概念をロボットに教示する。ロボットには真の参照点がどれなのかという情報は与えられない。各概念につき位置を変えながら10回ずつ教示を行った。学習の結果を図11に示す。図の原点が参照点の重心を表し、楕円は学習された各概念の位置分布（各概念が発話された時のロボットの位置を表す確率分布）である。Y軸のマイナス方向に話者がいることになるため、例えば「手前」の概念は参照点の重心よりも、話者側にロボットが位置することを表している。図から4つの概念が適切に学習できていることがわかる。



図7：相対位置概念の学習に用いたロボット



図8：地図画像を作成した環境



図9：作成した地図画像と参照点候補

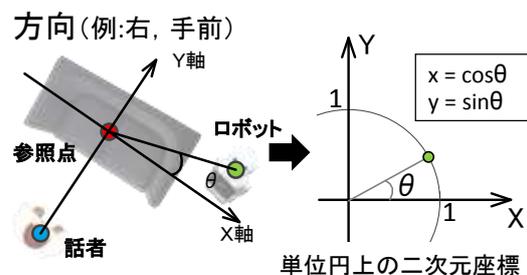
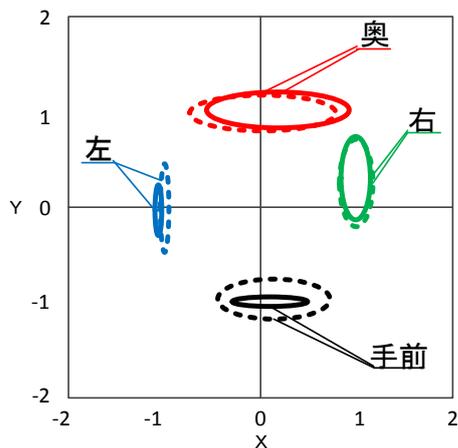


図10：学習に用いる「方向」の算出



方向の分布(中心 $\mu$ , 半径 $2\sigma$ )

--- 真の分布  
 —— 学習された分布

図 11 : 学習に用いる「方向」の算出

## 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)

[学会発表] (計 12 件)

- [1] 田口亮, DING WANG, QIYUE YU, 保黒政大, 梅崎太造, EM アルゴリズムを用いた絶対概念と相対概念の同時学習, 人工知能学会全国大会, 2G4-0S-19a-8, 2013
- [2] 田口亮, 東拓実, 保黒政大, 梅崎太造, ロボットによる発話理解のための相対的概念の学習, 計測自動制御学会システム・情報部門 学術講演会 (SSI2013), SS1-12, 2013
- [3] 沢登京介, 梅崎太造, 田口亮, 保黒政大, ヒューマンロボットコミュニケーションにおける意味学習機構の有用性, 情報処理学会 第 76 回全国大会, 4S-10, 2014
- [4] 東拓実, 加藤嗣, 田口亮, 梅崎太造, 保黒政大, 複数の座標系候補を持つ相対的な概念の学習, 情報処理学会 第 76 回全国大会, 5T-8, 2014 (学生奨励賞受賞)
- [5] 渡邊祐太, 田口亮, 服部公央亮, 保黒政大, 梅崎太造, サービスロボットによるインタラクションを通じた語彙の拡張, 人工知能学会全国大会, 2D4-3, 2015.
- [6] 藤本智也, 田口亮, 服部公央亮, 保黒政大, 梅崎太造, 自己組織化マップを用いた物体概念の学習, 人工知能学会全国大会, 2D4-5, 2015.
- [7] 渡邊祐太, 田口亮, 梅崎太造, 服部公央亮, 保黒政大, 言い直しを利用した語彙の拡張に関する検討, 電気・電子・情報関係学会東海支部連合大会 A1-1, 2015.
- [8] 藤本智也, 田口亮, 梅崎太造, 服部公央亮, 保黒政大, SOM と SVM を用いた姿勢変化に対応した物体認識, 電気・電子・情報関係学会東海支部連合大会 M1-8, 2015.
- [9] 丸山皓, 田口亮, 服部公央亮, 保黒政大, 梅崎太造, 距離センサを用いた地図画像からの参照点推定と相対的位置概念の学習, 計測自動制御学会 システム・情報部門 学術講演会, GS13-16, 2015.
- [10] 呉比, 田口亮, 服部公央亮, 保黒政大, 梅崎太造, 自動外観検査ロボットのための検査位置推定手法の検討, 計測自動制御学会 システム・情報部門 学術講演会, GS13-22, 2015.
- [11] 藤本智也, 渡邊祐太, 呉比, 田口亮, 服部公央亮, 保黒政大, 梅崎太造, キズ検査ロボットによる音声インタラクションを通じた語彙の拡張, 人工知能学会全国大会, 104-0S-22a-5, 2016 (発表予定).
- [12] Zhixiang Gu, Ryo Taguchi, Koosuke Hattori, Masahiro Hoguro, Taizo Umezaki, Learning of Relative Spatial Concepts from ambiguous instructions, The 13th IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design, and Evaluation of Human-Machine Systems, 2016. (載録決定)

[その他]

ホームページ

<http://taguchi-lab.com/>

## 6. 研究組織

(1) 研究代表者

田口 亮 (TAGUCHI, Ryo)

名古屋工業大学 大学院工学研究科 助教

研究者番号 : 70508415