

自律走行型ロボット ASKA による自動巡回システムの開発

東拓実^{*1}, 加藤嗣^{*1}, 服部公央亮^{*1}, 田口亮^{*1}, 保黒政大^{*2}, 梅崎太造^{*1}

^{*1}名古屋工業大学 ^{*2}中部大学

1. はじめに

家庭やオフィス等で人と対話するロボットは、その環境に固有の言語知識（人や物、場所の名前など）をユーザとのインタラクションを通して学習できなければならない。我々は、単語の知識を持たないロボットが、ユーザの多様な言い回しの発話から、単語の音素系列、および、単語と対象の間の直接的な対応関係（本稿ではこれを意味と呼ぶ）を学習するための手法を提案している[1, 2]。本稿ではこの手法をロボットによる自動巡回タスクに適用する。ロボットはまず、ユーザの誘導および教示から、建物内の地図と場所名を学習する。学習後、ロボットは獲得した単語知識を元に、ユーザの発話を認識し、指定された場所へ移動または巡回する。本稿ではそのためのシステムの枠組みについて説明する。

2. 語彙学習手法の概要

ユーザがある対象をロボットに提示し、音声でその名前を教示する。「これはボールペンです」等のように、教示には対象の名前以外の語を含む。本稿では、対象の名前を【キーワード】、キーワード以外の表現を【言い回し】と呼ぶ。言い回しとキーワードは独立であると仮定する。ロボットの初期知識は、各音素の音響モデルと、音素間の遷移モデルの二つだけである。教示された複数の音声-対象ペアから、単語の音素系列とその意味を学習する。

提案手法では、与えられた発話と指示対象の対応関係を、隠れ変数である単語列（単語ラベルの列）を介した共起確率モデルとして表現する。単語ラベルとその音素系列のペアは単語リストに記述する。単語リストを MDL 原理[3]に基づいて最適化することにより、発話と対象の対応関係を少ない単語数でうまくモデル化できるような単語集合を得ることができる。

2.1 発話と指示対象の共起確率モデル

発話 \mathbf{a} と対象を表す m 次元の連続ベクトル $\mathbf{o}=(o_1, o_2, \dots, o_m)^T$ の共起確率モデルを次式に示す。

Development of autopilot system with autonomous robot ASKA

Takumi Azuma^{*1}, Tsukasa Kato^{*1}, Koosuke Hattori^{*1}, Ryo Taguchi^{*1}, Masahiro Hoguro^{*2}, Taizo Umezaki^{*1},

^{*1} Nagoya Institute of Technology, ^{*2} Chubu University

$$\log P(\mathbf{a}, \mathbf{o}) \approx \max_s \{ \alpha \log P(\mathbf{a}|s) + \log P(s) + \log P(\mathbf{o}|s) \} \dots (1)$$

s は単語列である。 $P(\mathbf{a}|s)$ は音響モデルであり音素 HMM の連結として表現される。 $P(s)$ は文法モデルであり、単語 bigram として表現される。 $P(\mathbf{o}|s)$ は意味モデルであり次式で表す。

$$P(\mathbf{o}|s) = \prod_{i=1}^n \gamma(s, i) P(\mathbf{o}|w_i) \dots (2)$$

w_i は s に含まれる i 番目の単語、 $P(\mathbf{o}|w_i)$ は単語 w_i の意味、 $\gamma(s, i)$ は各単語の重み（単語の音素数より決定する）である。本稿では対象 \mathbf{o} を 2 次元の位置座標とし、 $P(\mathbf{o}|w)$ を 2 次元正規分布で表す。

2.2 語彙学習法

本学習は 4 つのステップに分けられる。 Step 1 は、学習データの全音声を音素列として認識し、その統計量から初期の単語リストを生成する。 Step 2 は、単語リストを用いて発話を単語列として認識し直し、その結果から意味モデルと文法モデルのパラメータを学習する。 Step 3 では学習結果を利用し単語リストの再構築を行う。ここでは、MDL 原理に基づき不要な単語の削除と、bigram 確率に基づいた単語連結を行う。 Step 4 では、再構築された単語リストを用いてモデルパラメータの再学習を行う。 Step 3, Step 4 を交互に繰り返すことで、より正しい音素系列が獲得される。手法の詳細は[1, 2]の通りである。

2.3 語彙学習法キーワードの出力

対象 \mathbf{o} が入力された場合、対応するキーワード w_0 を次式により出力する。

$$w_0 = \operatorname{argmax}_{w \in \Omega} \left\{ \log P(w) + \log P(\mathbf{o}|w) \right\} \dots (5)$$

Ω はキーワード集合である。キーワードの判定にはエントロピーの減少量 $I(w)$ を用いる。

$$I(w) = - \int P(\mathbf{o}) \log P(\mathbf{o}) d\mathbf{o} \dots (6) \\ + \int P(\mathbf{o}|w) \log P(\mathbf{o}|w) d\mathbf{o}$$

$I(w)$ が閾値以上の単語をキーワードと判定する。

2.4 語彙学習実験の結果

車輪移動型ロボット ASKA (図 1) を用いて場所名の学習を行った。ASKA はレーザレンジファインダを用いて地図作成と自己位置推定を行う。実験ではまず ASKA をリモコンで操作しながら建物内の地図を作成する。その後、ASKA を所定の場所へ移動させ計 10 種類の場所名を教



図 1: 車輪移動型ロボット ASKA

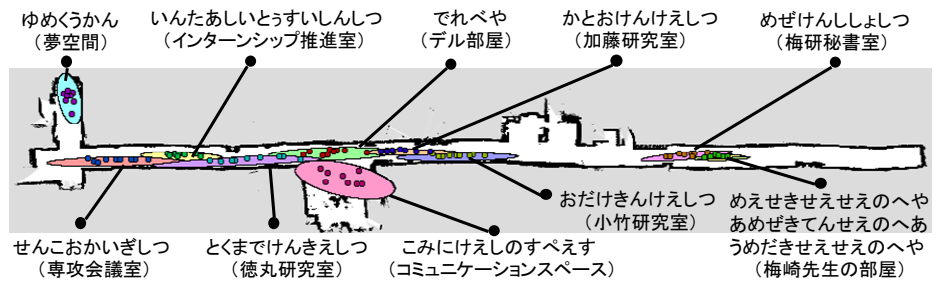


図 2: 教示したキーワード(括弧書き)と出力されたキーワード(かな表記)

示す。作成した地図と場所名を図 2 に示す。話者は男性 1 名で、教示する言い回し(表 1 参照)と位置を変えながら各場所で 9 回、計 90 回の教示を行った(図 2 の丸印)。なお、本実験では音声の収録と位置情報の取得は別々に行なった。3.2 の Step 3, Step 4 の繰り返しは 10 回とした。図 2 に獲得されたキーワードと、各キーワードの出力確率 $P(o,w)$ が閾値以上の部分を楕円で示す。この結果から、場所のカテゴリ化が行われていることがわかる。また、キーワードと判定された 12 単語の平均音素正解精度は 80% であった。学習なしで収録音声をもとに音素認識した際の音素正解精度は 76% であり正解精度の向上が見られた。これは単語リストの再構築時に音響的に有用な単語が取舍選択された結果である。

3. 獲得語彙を用いた対話型自動巡回システム

2 の手法により獲得された単語を用いることで、ユーザはロボットの巡回経路を発話で指示できるようになる。ロボットは「〇〇に行って」や「〇〇と××を巡回して」といった巡回を依頼する発話を認識するための文法知識を持っていると仮定する。ロボットは巡回依頼の発話を認識すると、発話中の場所名 w を取得し、 $P(o|w)$ が最大となる位置へ移動する。

図 3 にシステムのモジュール図を示す。音声認識には Julius, 音声合成には AITalk®を用いた。Julius の辞書には獲得された単語の音韻系列が登録されている。この音韻系列は音声合成にも使用できるため、対話中の移動先の確認等に使用できる。ASKA による巡回の例を図 4 に示す。この例では、ASKA を開始地点に設置し、開始点から「小竹研究室」を経由し「梅崎先生の部屋」までを移動した。図の青線は小竹研究室まで、赤線は「梅崎先生の部屋」までの自己位置推定結果である。ただし、開始地点の座標は既知とした。図から目的地に正しく移動できることがわかる。今後は、音声対話を利用した巡回指示の被験者実験を行い、提案手法の有効性を評価する。

4. まとめ

本稿では、多様な言い回しでの教示から、指示対象のカテゴリとそれを表す音素系列を同時に学習する手法を巡回ロボットに適用した。単語の学習では、10 種類の場所名を平均音素正解精度 80% で獲得できた。また、獲得した単語を用いた対話型の巡回システムを構築した。今回、巡回タスクでは認識できる言い回しを規定したが、今後は自由な発話に対して応答可能なシステムを目指し、キーワードスポッティングによる対話制御方法を検討する。また、語彙学習と巡回をシームレスに行う手法も検討する。

5. 謝辞

本研究の成果の一部は、科研費若手研究(B)(23700200)の助成を受けたものである。

参考文献

- [1] 田口 他: 人工知能学会論文誌, Vol.25, No.4, pp.5491-5501, (2010).
- [2] Taguchi, R., et al., Proc. of INTERSPEECH2011, pp.1325-1328. (2011)
- [3] Rissanen, J.: A universal prior for integers and estimation by minimum description length, The Annals of Stat., Vol. 11, No. 2, pp. 416-431, (1983)

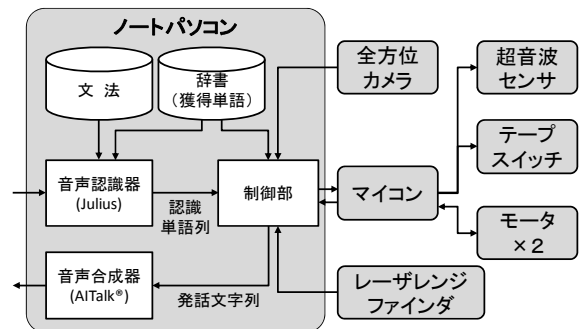


図 3: ASKA のモジュール図

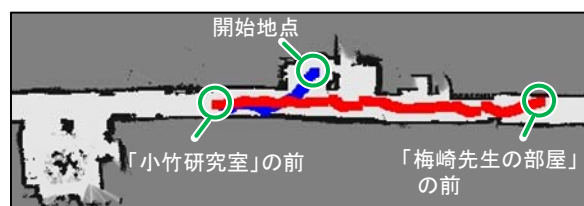


図 4: ASKA による移動結果