

不確実な音声認識を前提とした自己位置と語彙の同時推定モデル

Simultaneous estimation of self-location and lexicon based on ambiguous speech recognition

立命館大学 谷口彰, 吉崎陽紀, 谷口忠大

Akira Taniguchi, Haruki Yoshizaki and Tadahiro Taniguchi
Ritsumeikan University

Abstract In this paper, we proposed a novel learning method which can estimate self-location of a robot and names of places simultaneously. A robot performs a probabilistic self-localization from sensor data. We integrate ambiguous speech recognition results with it based on Bayesian approach. Experimental results show that a robot can obtain words for several locations and make use of them in self-localization task.

1 はじめに

近年, サービスロボット等の人間と共に人間の生活環境で動作するロボットの実現が求められている。人間の生活環境は様々であり, ロボットは未知な環境下でも周囲の環境を認知し学習していかなければならない。さらに, 人間の生活環境に存在するロボットは, 人とのインタラクションを通して知識を学習することが重要であると言える。また, ロボットはセンサを通じて環境の情報を認識するが, その際, 物理世界に存在する非常に多くの不確実性への対処が重要である [1]。

本稿では, 場所についての知識獲得に注目し, 自己位置推定を行いながら環境を移動するロボットに, 人が場所の名前を教示することで場所に対応した語彙を獲得させることを目標とする。ロボットが自己位置推定を行う際にはセンサ情報から確率的な自己位置推定を行うが, 本稿ではこれに不確実な音声認識結果を統合することを行う。すなわち本稿では, 不確実な音声認識と自己位置情報を相互に有効活用することで, 自己位置と語彙の同時推定モデルを提案する。

山田ら [2] の研究では, 対象の名前以外の語を含む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が提案されている。学習後, 学習に用いていない位置情報を入力し, 場所に対応した正しいキーワードの出力を行っているが, 本研究では獲得した知識からロボット自身の位置情報へのフィードバックを行うモデルを提案する。

2 自己位置と語彙の推定モデル

本研究では, 自己位置推定の手法には MCL (Monte Carlo Localization) を利用し, MCL に位置概念を導入し拡張したモデルを提案する。位置概念とは, 場所に対

応した語彙のことであり, これには場所の名前とその名前によって表現される場所の位置分布という二つの知識が含まれるものとする。

移動ロボットに地図のある環境上で建物内を移動させて, 自己位置推定を行わせることを想定する。ロボットを学習対象の場所に移動させ, 人が場所の名前を複数回教示することで, その場所に対応した語彙を学習させる。提案手法の全体像を表す概略図を図 1¹に示す。

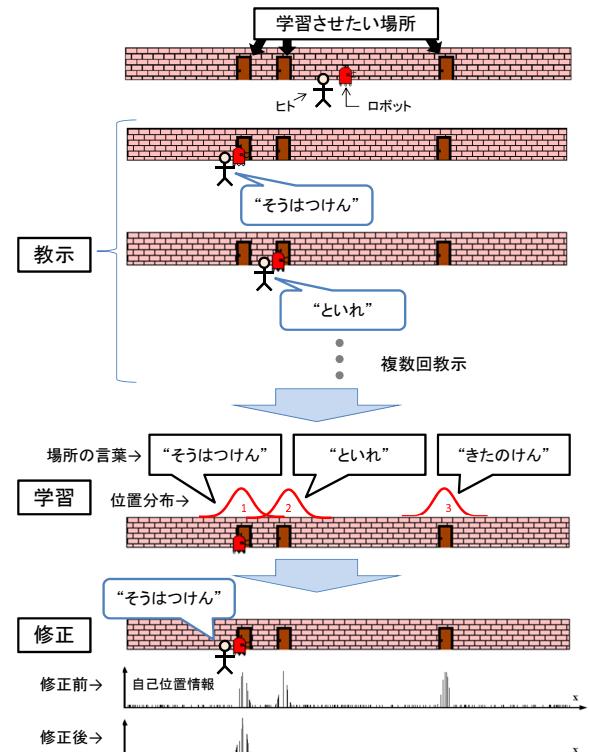


図 1: 提案手法の全体像を表す概略図

¹図の引用元(一部): Probabilistic Robotics: <http://probabilisticrobotics.com/> 取得日:2012/07/01

ロボットは、事前に単語の知識を持たず、日本語の音節のみを認識できるものとする。

本稿では、位置概念を獲得したロボットが場所の名前を聞いたときの自己位置情報の修正方法と、場所の名前を複数回教示されたときの位置概念の学習方法について考える。

2.1 位置概念の定義

位置概念は、場所の名前と場所の位置分布を含むものである。学習により獲得した位置概念の集合を λ とし、式(1)の様に表す。 L は位置概念の数である。 i 番目の位置概念は、式(2)の様に定義する。

$$\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_L\} \quad (1)$$

$$\lambda_i = \{W_i, \mu_i, \Sigma_i\} \quad (2)$$

このとき、 W_i は学習により得た i 番目の場所の名前である。式(3)の l_j^i は W_i の j 番目の文字であり、 d_i は W_i の文字数を表す。

$$W_i = (l_1^i l_2^i \cdots l_{d_i}^i) \quad (3)$$

場所の名前は、式(4)の様に表せる。また、場所の位置分布はガウス分布を仮定する。そのため、 μ_i は i 番目の分布の平均、 Σ_i は i 番目の分布の分散である。平均は、式(5)、分散は、式(6)の様に表せる。

$$\mathbf{W} = \{W_1, W_2, \dots, W_L\} \quad (4)$$

$$\boldsymbol{\mu} = \{\mu_1, \mu_2, \dots, \mu_L\} \quad (5)$$

$$\boldsymbol{\Sigma} = \{\Sigma_1, \Sigma_2, \dots, \Sigma_L\} \quad (6)$$

時刻 t において、ロボットが聞き取った音声 O_t に対応した位置概念の指標を C_t で表す。 C_t は、1 から L までの離散値をとる。位置概念の指標の集合を C とし、式(7)に示す。

$$C_t \in \mathbf{C} = \{1, 2, \dots, L\} \quad (7)$$

つまり、位置概念 λ_{C_t} は、 $\lambda_{C_t} = \{W_{C_t}, \mu_{C_t}, \Sigma_{C_t}\}$ と表せ、 W_{C_t} μ_{C_t} Σ_{C_t} は、それぞれ C_t 番目の位置概念の場所の名前、位置分布の平均、分散を表す。

2.2 位置概念を導入した自己位置推定

MCL に位置概念の指標 C_t 、場所の位置分布 $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ と場所の名前 \mathbf{W} 及び音声認識した単語 O_t の関係を加えたグラフィカルモデルを図2に示す。青い矢印で示した部分が従来の MCL、赤の矢印で示した部分が、提案手法による拡張部分である。グラフィカルモデルの各要素

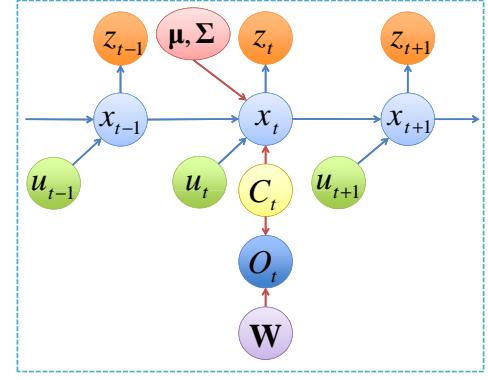


図2: 位置概念 MCL のグラフィカルモデル

表1: グラフィカルモデルの各要素表

x_t	ロボットの自己位置
u_t	制御値
z_t	計測値
C_t	位置概念の指標
O_t	音声認識した単語
\mathbf{W}	学習した場所の名前
$\boldsymbol{\mu}, \boldsymbol{\Sigma}$	場所の位置分布 (平均, 分散)

についてまとめたものを表1に示す。

MCL の導出式に O_t を加えたものを、式(8)に示す。

$$\begin{aligned} p(x_{0:t} | z_{1:t}, u_{1:t}, O_{1:t}) \\ \propto p(z_t | x_t) p(O_t | x_t) p(x_t | x_{t-1}, u_t) \\ \cdot p(x_{0:t-1} | z_{1:t-1}, u_{1:t-1}, O_{1:t-1}) \end{aligned} \quad (8)$$

式(8)で導出された $p(O_t | x_t)$ の展開を、式(9)に示す。これは、 x_t という場所で O_t という言葉を聞いたときの確率を表しており、計測モデル $p(z_t | x_t)$ と同じく尤度として計算される。式(9)の3行目は、グラフィカルモデルの関係により導出される。

$$\begin{aligned} p(O_t | x_t) &= \sum_{C_t} p(O_t | C_t) p(C_t | x_t) \\ &\propto \sum_{C_t} p(O_t | C_t) p(x_t | C_t) p(C_t) \\ &= \sum_{C_t} p(O_t | \mathbf{W}, C_t) p(x_t | \boldsymbol{\mu}, \boldsymbol{\Sigma}, C_t) \frac{1}{L} \end{aligned} \quad (9)$$

本研究では、 $p(C_t)$ は無情報を仮定し、一様分布として $p(C_t) = \frac{1}{L}$ の様に近似する。

$p(O_t | \mathbf{W}, C_t)$ の計算については、編集距離を使用し場所の名前と認識文字の近さを表す確率の式を、式(10)

表 2: 位置概念学習の処理の流れ

- (1) μ, Σ を初期化
- (2) \mathbf{W} の情報無しで, C_t をサンプリング
- (3) μ, Σ をサンプリング
- (4) \mathbf{W} をサンプリング
- (5) \mathbf{W} の情報有りで, C_t をサンプリング
- (6) (3) ~ (5) の過程を複数回繰り返す

に定義する。LD は編集距離を計算する関数であり, β は LD の値による影響度合いを確率式に反映するパラメータである。

編集距離とは, 二つの文字列の異なり具合を量量化するもので, 文字の挿入・削除・置換で一方を他方に変形するための最小手順回数を数えたものである。本研究では, 音声認識器による認識単語や学習した場所の名前の音の近さ, 言葉の近さを編集距離によって測る。

$$p(O_t | \mathbf{W}, C_t) \propto \exp(-\beta LD(O_t, W_{C_t})) \quad (10)$$

$p(x_t | \mu, \Sigma, C_t)$ の計算については, 場所の位置分布と自己位置情報の近さを表す確率の式を, 式 (11) に定義する。この式は, 多変量ガウス分布を表している。 D は次元数である。

$$\begin{aligned} p(x_t | \mu, \Sigma, C_t) \\ \propto \frac{1}{(2\pi)^{\frac{D}{2}} |\Sigma_{C_t}|^{\frac{1}{2}}} \exp \left(-\frac{1}{2} (x_t - \mu_{C_t})^T \Sigma_{C_t}^{-1} (x_t - \mu_{C_t}) \right) \end{aligned} \quad (11)$$

2.3 位置概念の場所の位置分布と場所の名前の推定

オフライン学習により, 発話場所と認識単語による複数の学習用データから場所の位置分布と場所の名前をギブスサンプリングによって推定する。位置概念の数 L は既知とする。ギブスサンプリングによる位置概念学習の処理の流れを表 2 に, そのイメージ図を図 3 に示す。三つの青い四角が学習対象の場所を表している。

各ステップについて説明する。

1. μ, Σ を初期化

位置概念の初期分布は全て, $\mu_c =$ (地図の範囲内に一様乱数) $\Sigma_c = \begin{bmatrix} \sigma_{initial} & 0 \\ 0 & \sigma_{initial} \end{bmatrix}$ とする。

2. \mathbf{W} の情報無しで, C_t をサンプリング

このステップでは, どのデータがどの位置概念の指標を表すかを推測する。 C_{T_o} 以外の他の全ての

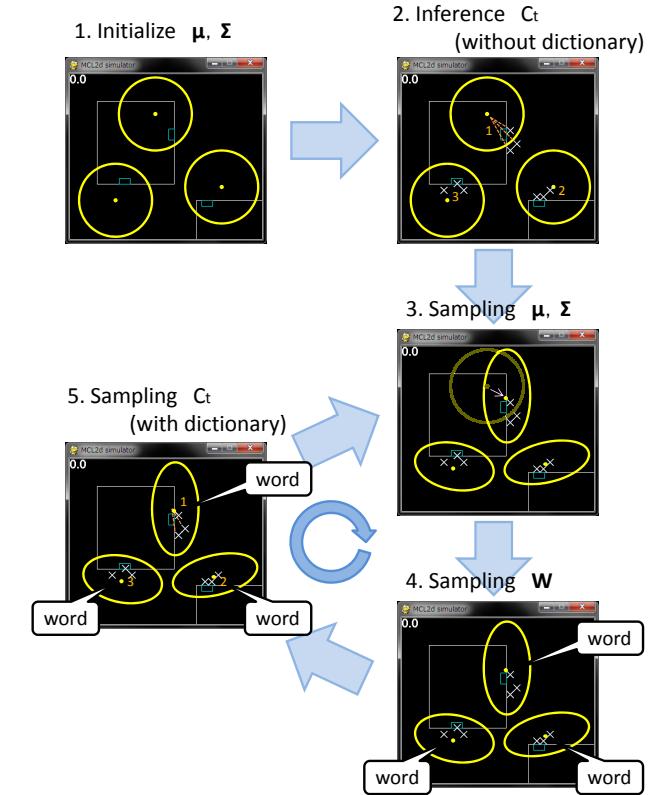


図 3: ギブスサンプリングによる位置概念学習の流れ

変数が与えられた上で C_{T_o} の確率を表す式を, 式 (12) に示す。

$$\begin{aligned} p(C_{T_o} | \mathbf{W}, O_{T_o}, x_{T_o}, \mu, \Sigma) \\ = \prod_{t \in T_o} p(C_t | \mathbf{W}, O_t, x_t, \mu, \Sigma) \\ \propto \prod_{t \in T_o} p(x_t | \mu_{C_t}, \Sigma_{C_t}) p(O_t | C_t, \mathbf{W}) p(C_t) \end{aligned} \quad (12)$$

これを, 式 (13) の様に $t \in T_o$ についてそれぞれサンプリングする。

$$\begin{aligned} C_t &\sim p(C_t | \mathbf{W}, O_t, x_t, \mu, \Sigma) \\ &\propto p(x_t | \mu_{C_t}, \Sigma_{C_t}) p(O_t | W_{C_t}) \end{aligned} \quad (13)$$

C_t のサンプリングは, $C_t \in \mathbf{C} = \{1, 2, \dots, L\}$ についてそれぞれ確率を出し, 亂数により選択する。(2) ステップでは, \mathbf{W} の情報が推定されていないため, 式 (14) を計算し, サンプリングする。

$$C_t \sim p(x_t | \mu_{C_t}, \Sigma_{C_t}) \quad (14)$$

3. μ, Σ をサンプリング

場所の位置分布の学習の式を，式 (15) に示す .

$$\begin{aligned} p(\mu, \Sigma | \mathbf{W}, O_{T_o}, x_{T_o}, C_{T_o}) \\ = \prod_{c \in \mathbf{C}} \left(\prod_{\substack{c=C_t \\ t \in T_o}} p(x_t | \mu_c, \Sigma_c) p(\mu_c, \Sigma_c) \right) \end{aligned} \quad (15)$$

これを，式 (16) の様に $c \in \mathbf{C} = \{1, 2, \dots, L\}$ についてそれぞれ計算し，サンプリングする .

$$\begin{aligned} \mu_c, \Sigma_c &\sim p(\mu_c, \Sigma_c | \mathbf{W}, O_{T_o}, x_{T_o}, C_{T_o}) \\ &= \prod_{\substack{c=C_t \\ t \in T_o}} \left(p(x_t | \mu_c, \Sigma_c) \right) p(\mu_c, \Sigma_c) \end{aligned} \quad (16)$$

これは，学習データの中 ($t \in T_o$) で $C_t = c$ のデータのみを選んで， $p(\mu_c, \Sigma_c)$ に掛け合わせている . また， $\prod_{\substack{c=C_t \\ t \in T_o}} p(x_t | \mu_c, \Sigma_c)$ を尤度関数と見たとき， $p(\mu_c, \Sigma_c)$ が事前分布となり，式 (17) の形で表せる . Λ は精度行列であり，共分散行列の逆行列である ($\Lambda = \Sigma^{-1}$).

$$p(\mu, \Lambda | X) \propto p(X | \mu, \Lambda) \times p(\mu, \Lambda) \quad (17)$$

これを式 (18) の様に，事前分布にガウス-ウィシャート分布を用い，ガウス分布に対するベイズ推論 [3] を行うことにより事後分布を求める . m_0, β_0, W_0, ν_0 は，ハイパーパラメータである .

$$p(\mu, \Lambda) = \mathcal{N}(\mu | m_0, (\beta_0 \Lambda)^{-1}) \mathcal{W}(\Lambda | W_0, \nu_0) \quad (18)$$

つまり， μ, Λ のサンプリングは以下の様に行う . まず，式 (19) の様に， Λ をウィシャート分布からサンプリングする . 次に，式 (20) の様に，サンプリングした Λ と m_0, β_0 を用いて μ をサンプリングする .

$$\Lambda \sim \mathcal{W}(\Lambda | W_N, \nu_N) \quad (19)$$

$$\mu \sim \mathcal{N}(\mu | m_N, (\beta_N \Lambda)^{-1}) \quad (20)$$

学習データの中に該当する位置概念がない場合，パラメータ m_0 の値を地図の範囲内となるような一様乱数により与える .

4. \mathbf{W} をサンプリング

場所の名前の学習の式を，式 (21) に示す .

$$\begin{aligned} p(\mathbf{W} | \mu, \Sigma, O_{T_o}, x_{T_o}, C_{T_o}) \\ = \prod_{c \in \mathbf{C}} \left(\prod_{\substack{c=C_t \\ t \in T_o}} p(O_t | W_c) p(W_c) \right) \end{aligned} \quad (21)$$

これを，式 (22) の様に $c \in \mathbf{C} = \{1, 2, \dots, L\}$ についてそれぞれ計算し，サンプリングする .

$$W_c \sim \prod_{\substack{c=C_t \\ t \in T_o}} \left(p(O_t | W_c) \right) p(W_c) \quad (22)$$

このとき， $p(W_c)$ は，無情報を仮定し，一様分布として $p(W_c) = \frac{1}{N}$ の様に近似する .

W_c のサンプリングは， $t \in T_o = \{t_1, t_2, \dots, t_N\}$ についてそれぞれ確率を出し，乱数により選択する . 学習データの中に該当する位置概念がない場合， $p(W_c)$ 一様分布より，全ての学習用データの認識単語から等確率で選択する .

5. \mathbf{W} の情報有りで， C_t をサンプリング

(2) ステップとは違い， \mathbf{W} の情報が推定されているため，式 (13) を計算し， C_t をサンプリングする . サンプリング方法は，(2) と同様である .

6. (3) ~ (5) の過程を複数回繰り返す .

3 実験

人が場所の名前を発話したときの音声情報の認識には，大語彙連続音声認識システム Julius[4]² を利用する .

Julius の単語辞書には，既存の大量語が登録された単語辞書を使わず，日本語音節のみを登録した単語辞書を使用する . これにより，発話音声を音節ごとに分けて音節列として認識することができる . 実験に使用したマイクは，SHURE 社の PG27 USB³ である .

3.1 場所の名前の教示による自己位置情報の修正

3.1.1 実験条件

1 次元空間上での MCL で実験を行った . ロボットは画面内を左右に移動する . ロボットの左右にはセンサがあり，センサ限界値以内にドアが存在する場合，ドアまでの距離を返す . 三つのピンクの四角がドアを表し，ロボットの下部に分布しているピンクの丸がパーティクル，その下の青い線がパーティクルに対応した尤度を表している . ロボットの初期座標は (50, 50)，パーティクル数は 500 個で画面内 (x 座標 : 0 ~ 600) に一様に分布，センサ限界値は左右共に 50pixel とした .

三つのドア付近で既に位置概念を適切に獲得しているものとし，場所の名前と位置分布を事前に設定する . 位

²インストールバージョンは dictation-kit-v4.2-win である .

³PG27USB サイドアドレス型コンデンサー・マイクロホン SHURE 社: <http://www.shure.co.jp/ja/products/microphones/pg27usb>

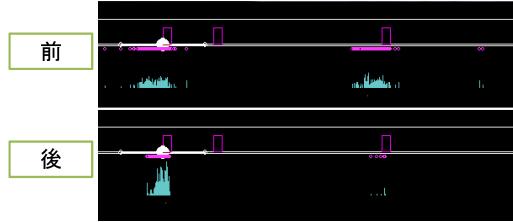


図 4: 言語情報を与えた前後でのパーティクルの変化

置概念の数を $L = 3$, 場所の名前を $\mathbf{W} = \{“\text{そうはつけん}”, “\text{といれ}”, “\text{きたのけん}\}\}$, 場所の位置分布を $\mu = \{125, 185, 385\}$, $\Sigma = \{10, 10, 10\}$ と設定した .

本実験では , 位置概念のある場所付近で , 必ずその位置概念を表す言葉 W_{C_t} が発話されるものとした .

3.1.2 実験結果

例として , 一番左のドアで行った場合の結果を示す . 初期状態からロボットを右へ動かし続けると , パーティクルが一番左のドアと一番右のドア周辺に多く分布するような状況が起こり得る . その様子が図 4 の上部である . ロボットは実際には一番左のドア付近にいるため , 一番左のドアの名前“ そうはつけん ”を言語情報として教示する . 言語情報を教示したときの実行画面の一部が図 4 の下部である . 一番左のドア付近に存在するパーティクルの尤度がより高くなり , ロボットの真の位置に近いパーティクルが増えるという結果になった . また , 他のドアについても同様に言語情報を教示すると , ロボットの真の位置に近いパーティクルが増えるという結果が得られた . この結果から , 言語情報を教示することによって自己位置情報の修正が行えていると考えられる .

3.2 位置概念の学習

3.2.1 実験条件

座標原点は左上とし , x 軸は右方向 , y 軸は下方向の 2 次元空間上で実験を行った . 本実験では , $\sigma_{initial} = 10000$, (3) ステップにおける事前分布のハイパーパラメータの初期値は , $\beta_0 = 0.001$, $m_0 = [0, 0]^T$, $W_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\nu_0 = 1$ とし (6) ステップでの繰り返し回数は , 10 回とした . 位置概念の数は $L = 3$ とした . 発話場所は , それぞれ三つのドア前付近とし , 10 回ずつ , 合計 30 回分のデータを毎回 , 各ドア前の座標にガウスノイズを付加し生成した . 各ドア前の座標は , $doorA = (100, 210)$, $doorB = (200, 110)$, $doorC = (250, 220)$ である . 各ドア前に付加したガウスノイズの平均は , $[0, 0]^T$ である .

発話単語	認識単語				
	ぶはつけん	そふつきゅ	そつけん	すがつけん	そぶはつけん
そうはつけん	ぶはつけん	そふつきゅ	そつけん	ぶはつけん	そぶはつけん
しょむはつけん	しょむはつけん	そつけん	そつけん	ぶはつけん	そほつけん
ちのうけん	ちのけん	しのけん	きのおけん	ちのうけん	ちのけん
ちのけん	ちのけん	ちのけん	ちのけん	ちのけん	ちのけん
かいだん	かりだん	かいだん	かいだん	かにたん	かにだん
かいだん	かいだん	かいだん	かいだん	かいるん	かいだん

図 5: 日本語音節単語辞書を用いた Julius の認識結果

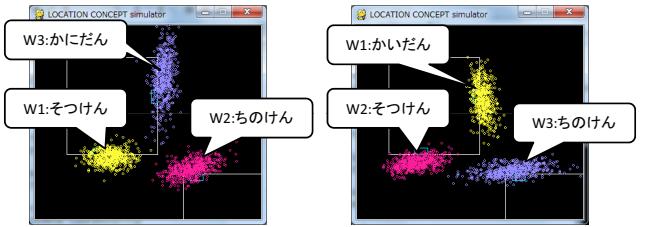


図 6: 学習結果例 1

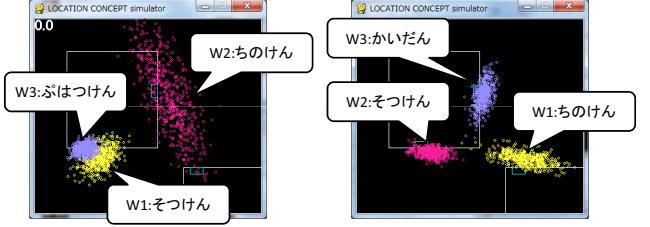


図 7: 学習結果例 2

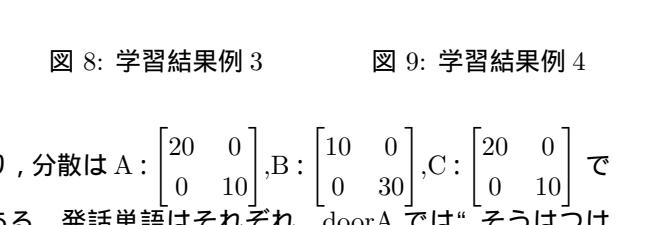


図 8: 学習結果例 3

り , 分散は $A : \begin{bmatrix} 20 & 0 \\ 0 & 10 \end{bmatrix}$, $B : \begin{bmatrix} 10 & 0 \\ 0 & 30 \end{bmatrix}$, $C : \begin{bmatrix} 20 & 0 \\ 0 & 10 \end{bmatrix}$ である . 発話単語はそれぞれ , doorA では“ そうはつけん ” , doorB では“ かいだん ” , doorC では“ ちのうけん ”とし , 学習用データには , 図 5 の各発話単語に対する認識結果を利用した .

3.2.2 実験結果

この設定で試行した 15 回分の学習結果から , 4 例を図示したものを図 6 , 図 7 , 図 8 , 図 9 に示す . 各色の点群は , 学習した場所の位置分布に従う点を各位置概念に対して 500 個ずつ描画したものである . W1 ~ W3 は , それぞれの分布に対応した場所の名前である . 今回は , 三つのドア前に対応した三つの位置概念を学習させることを行った . 15 回分の学習結果中 , 2 回のみで図 8 の様に明らかに , 各ドアに対して正しく 1 対 1 対応で学習できていない結果が見られた . その他の学習結果では , 図 6 , 図 7 , 図 9 の様な学習結果が得られ , 多くの場合で三つのドア前に位置概念が学習できていることがわかる .

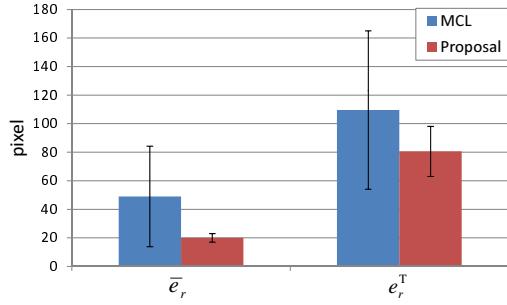


図 10: MCL と提案手法による推定誤差の比較

3.3 学習した位置概念の評価のための比較実験

3.3.1 実験条件

次に、先の位置概念の学習結果を用いた自己位置情報の修正について確認した。学習時と同じ 2 次元空間上で通常の MCL と提案手法によって位置概念を学習した状態の MCL とで自己位置の推定誤差精度の比較を行った。ロボットは 300 step 移動し、各ドア前を“かいだん”、“ちのうけん”、“そはつけん”と順番に巡る様な動作をする。この時、全ての試行において、ロボットは同じ動作とした。提案手法でのみ、それぞれドア前付近に来たとき場所の名前を教示した。教示単語は、図 5 の各ドアに対する認識単語からランダムで選ぶ。パーティクル数は $M = 300$ とし、ロボットは 1 step 毎に 1 回の動作を行う。MCL における推定誤差の評価には、上田ら [5] の評価方法を参考にした。まず、step 毎に xy 平面での推定誤差を表す、式 (23) を記録する。

$$e_t = \sqrt{(\bar{x}_t - x_t^*)^2 + (\bar{y}_t - y_t^*)^2} \quad (23)$$

$\bar{x}_t = \sum_i^M w_t^{(i)} x_t^{(i)}$, $\bar{y}_t = \sum_i^M w_t^{(i)} y_t^{(i)}$ であり、 x_t^*, y_t^* はロボットの真の位置座標である。試行後、 e_t の平均値 \bar{e}_r と、区間 $[0, \gamma]$ が e_t の 95% 以上を含む γ の最小値 e_r^T を計算する。 e_r^T は推定の安定性を示す指標として見せる。

3.3.2 実験結果

15 回分の実験結果を \bar{e}_r , e_r^T によって比較したグラフを、図 10 に示す。どちらの指標でも、推定誤差は提案手法の方が小さいという結果が得られた。また、この結果で有意水準 5% の片側 t 検定を行ったところ、 \bar{e}_r , e_r^T 共に有意差が見られた。よって、学習により、自己位置情報の修正を行えるような位置概念が獲得できていることが確認できた。

4 まとめと考察

本研究では、移動ロボットを使用し場所に対応した語彙の獲得を目的とし、位置概念を獲得する方法と自己位置情報を修正する方法を検討した。位置概念の学習では、多くの場合で三か所あるドアの前に位置概念がそれぞれ形成されたが、一つのドアの前に二つの位置概念が形成され、残りの二つのドアを一つの位置概念が包含するような分布が形成された場合も見られた。

今後の発展として、位置概念の数 L を既知から未知にも対応することや、オンライン学習への対応などが考えられる。また本研究では、MCL における地図情報がある場合でのモデルであったが、SLAM (Simultaneous Localization And Mapping) を使用した地図情報が与えられていない場合への拡張が考えられる。

本稿では、簡易なシミュレータ上での実験を行ったが、今後は SIGVerse[6] を用いたより現実的な系での実験を行う予定である。

参考文献

- [1] S. Thrun, W. Burgard, D. Fox, 上田隆一 (訳). 確率ロボティクス. 毎日コミュニケーションズ, 2007.
- [2] 山田雄治, 服部公央亮, 田口亮, 梅崎太造, 保黒政大, 岩橋直人, 船越孝太郎, 中野幹生. 連続音声から場所の名前を学習する自律移動ロボット. 一般社団法人情報処理学会 全国大会講演論文集, Vol. 2011, No. 1, pp. 237–239, mar 2011.
- [3] C.M. ビショップ著 元田 浩・栗田多喜夫 監訳 樋口知之他監訳. パターン認識と機械学習 上 ベイズ理論による統計的予測. 丸善出版, 2007.
- [4] 荒木雅弘. フリーソフトでつくる音声認識システム: パターン認識・機械学習の初步から対話システムまで. 森北出版株式会社, 2007.
- [5] 上田隆一, 新井民夫, 浅沼和範, 梅田和昇, 大隅久. パーティクルフィルタを利用した自己位置推定に生じる致命的な推定誤りからの回復法. 日本ロボット学会誌, Vol. 23, No. 4, pp. 466–473, 2005.
- [6] 稲邑哲也, 柴田智広, 瀬名秀明, 橋本敬, 川合伸幸, 宮下敬宏, 櫻井圭記, 清水正宏, 大武美保子, 細田耕, 梅田聰, 乾健太郎. 2a2-c22 社会的インタラクションを実装可能なシミュレータープラットフォーム: Sigverse:社会的知能発生シミュレータ. ロボティクス・メカトロニクス講演会講演概要集, Vol. 2009, pp. 2A2-C22(1)–2A2-C22(4), 2009.