

# EM アルゴリズムを用いた参照点に依存した語意の学習

田口 亮<sup>\*1</sup>, WANG DING<sup>\*1</sup>, YU QIYUE<sup>\*1</sup>, 保黒政大<sup>\*2</sup>, 梅崎太造<sup>\*1</sup>

<sup>\*1</sup>名古屋工業大学, <sup>\*2</sup>中部大学

## 1. はじめに

近年, 単語の意味をロボットの感覚・運動系の信号と対応付けて学習させる研究が行われ始めた[1, 2]. 従来研究の多くは, 単語と共に起るセンサ情報を直接的に学習していた. 例えば, 文献[2]では画像から抽出された物体の色や形, 大きさなどの特徴量と音声の対応関係を学習している. しかし, 色や大きさ, 位置などの概念は, 他の物体との比較や, 典型的な概念との比較によって表される. 例えば, 蟻を指さして「小さい」と言う場合もあれば, 同じ蟻でも隣の蟻と比較して「大きい」と言う場合もある. また, 「小さい象」は通常, 「大きい蟻」よりも大きい. このように言葉には参照点に依存し, 相対的な意味を持つ単語が多く存在する.

先の例では「大きい蟻」や「小さい象」などのように, 参照点が言語的に明示されていたが, 人間同士の対話においては, 参照点が明示されない場合がある. 例えば, 図 1(a)のような場面で, 人が「右にあるやつを取って」と発話した場合, 取って欲しいものは画像右側のぬいぐるみであることが想像できる. この場合, 参照点は画像の中心, または話者の視線の中心となる. 一方で図 1(b)のような場面で, 「左にあるやつを取って」と発話することも可能である. この場合, 取って欲しいものは画像中央の赤い箱であり, 参照点は画像右側の白い箱になる. 参照点に依存した語意を学習する際も同様に, 参照点が明示されない場合でも, 複数の候補から適切に参照点を推定しながら語意を学習していく必要がある. 本稿では EM アルゴリズム[3]を用いることで, 参照点の推定と語意の学習とを同時に行う手法を提案する.

## 2. 語彙学習手法

ユーザがある対象をロボットに提示し, 対象が持つ特徴の名前を教示する. 本実験では, 音声の代わりにテキストのラベルを与える. ここで, 教示ラベルを  $w$ , 対象の特徴を  $x_n$  とする.  $n$  は学習サンプルのインデックス ( $1 \leq n \leq N$ ,  $N$



図 1: 参照点が明示されない例

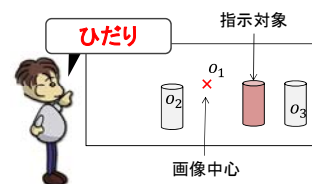


図 2: 教示の例 ( $o_1 \sim o_3$  が参照点候補)

は学習サンプル数) である. また概念の比較対象となる参照点の候補を  $k$  とする. 参照点  $k$  を基準とした特徴量を  $x_{nk}$  と表す. 各場面における指示対象と参照点の候補は既知とするが, 候補のいずれかが真の参照点なのかは未知である. 例えば, 特徴量  $x_n$  が指示対象の位置を表す場合,  $x_{nk}$  は参照点  $k$  からの相対位置を表す. 各学習サンプルでは, 参照点候補の数に応じて, 学習候補となる特徴量  $x_{nk}$  が複数現れる. 図 2 の例では, 参照点の候補として  $o_1 \sim o_3$  が与えられ, 各参照点候補と指示対象との距離が学習候補の特徴量  $x_{n1} \sim x_{n3}$  となる. 図では  $o_1, o_2$  の右に指示対象があるため, 「ひだり」の意味を学習する際には,  $o_3$  を参照点とし, 他の特徴量候補は破棄しなければならない. このように, 本タスクでは, 学習サンプル毎に正解となる特徴量のみを採用し, 学習していくことが望ましい.

本稿では, 与えられたラベルと指示対象の対応関係を平均  $\mu$ , 分散  $\sigma^2$  の正規分布  $N(x_{nk}|\mu, \sigma^2)$  でモデル化する. また, 学習サンプル  $n$  において参照点候補  $k$  が真の参照点である確率を  $\pi_{nk}$  とおく. そして式(1)に示すモデル尤度を最大化するパラメータ  $\theta = (\mu, \sigma^2, \pi_{nk})$  を EM アルゴリズムにより求める. 各パラメータの更新式を式(2)~(4)に示す.

$$\theta = \arg \max \sum_{n=1}^N \ln \sum_{k=1}^{M_n} \pi_{nk} N(x_{nk}|\mu, \sigma^2) \quad \dots(1)$$

$$\mu = \frac{1}{N} \sum_{n=1}^N \frac{\sum_k \pi_{nk} N(x_{nk}|\mu, \sigma^2) x_{nk}}{\sum_k \pi_{nk} N(x_{nk}|\mu, \sigma^2)} \quad \dots(2)$$

Reference-Point-Dependent Lexical Learning Method Using EM Algorithm

Ryo Taguchi<sup>\*1</sup>, WANG DING<sup>\*1</sup>, YU QIYUE<sup>\*1</sup>, Masahiro Hoguro<sup>\*2</sup>, Taizo Umezaki<sup>\*1</sup>,

<sup>\*1</sup>Nagoya Institute of Technology, <sup>\*2</sup>Chubu University

$$\sigma^2 = \frac{1}{N} \sum_{n=1}^N \frac{\sum_k \pi_{nk} N(x_{nk} | \mu, \sigma^2) (x_{nk} - \mu)^2}{\sum_k \pi_{nk} N(x_{nk} | \mu, \sigma^2)} \quad \dots(3)$$

$$\pi_{nk} = \frac{\pi_{nk} N(x_{nk} | \mu, \sigma^2)}{\sum_j \pi_{nj} N(x_{nk} | \mu, \sigma^2)} \quad \dots(4)$$

EM アルゴリズムは隠れ変数を含む確率モデルのパラメータを、最尤法に基づいて推定する手法であり、混合正規分布の学習にも用いられる。本手法ではサンプル毎に参照点の確率 $\pi_{nk}$ を定義しているが、混合正規分布では各サンプルに対して独立な混合重み $\pi_k$ を用いる。また、混合正規分布の場合には複数の正規分布を用いるが、本手法では単一の正規分布となっている。

また EM アルゴリズムを用いた動作概念の獲得に関する研究も行われている[4]。この研究では参照点を基準とした軌跡で動作をモデル化している。本研究はこの研究を動作以外の概念に拡張したものである。

### 3. 実験結果と考察

#### 3.1. 位置を表す単語の学習

図1のような環境を用いて「上、下、左、右」の4単語の学習実験を実施した。画像データは1単語につき10枚であり、1つの画像には1~4つの物体を含む。特徴量は指示対象の位置座標(x, y)とした。参照点候補は画面中心および他の物体の重心位置である。「左、右」の学習結果を図3に示す。提案手法を用いずに全ての特徴量候補を用いると、適切に学習できないことがわかる。一方、提案手法ではX軸において「左、右」の分布が分離されており、相対的な意味の学習ができたと言える。また、「上、下」についても同様に学習できることが確認された。

#### 3.2. 色を表す単語の学習

「赤、青、黄、緑、灰色、明るい、暗い」の7単語の学習実験を実施した。実験には google 画像検索により得られた画像を使用した。検索の際のキーワードは、「Dark yellow」など、明るさを表す単語と色を表す単語をペアにした。画像データは1単語につき20枚である。使用した特徴量は、画像から RGB の最頻値を抽出し、その値を色相、彩度、明度を表す HSV 表色系に変換したものである。本実験ではまず「赤、青、黄、緑、灰色」の基本色を参照点なしで学習した後、「明るい、暗い」の単語を学習した。その際に指示対象となる色画像の隣に、ランダムで選んだ色画像を並べて提示する。指示対象が明るい暗いかは、隣接した画像と比較するか、指示対象の色の典型例と比較して判断する。すなわ

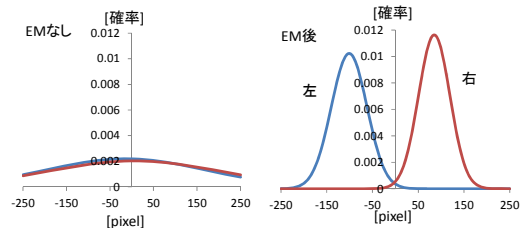


図3: 「左」「右」の学習結果

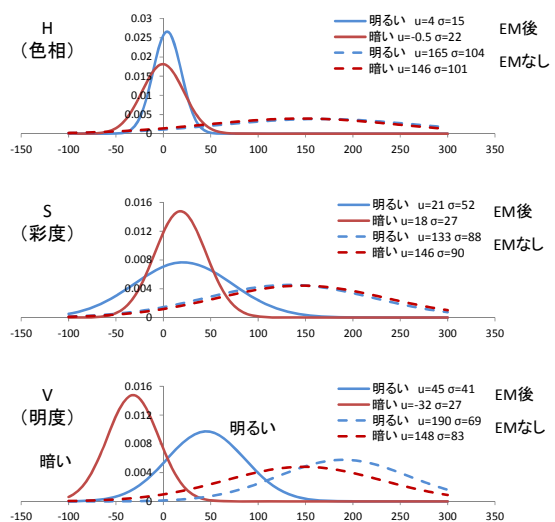


図4: 「明るい」「暗い」の学習結果

ち参照点は2つである。色の典型例としては事前に学習した5単語の平均色で、指示対象と同じ色相のものを用いた。学習の結果を図4に示す。提案手法を用いない場合には、色相、彩度、明度共に分布の分離は見られないが、EM アルゴリズムで学習することにより、明度の軸で「明るい」と「暗い」の分布が分離し、正しく学習できていることが確認された。

### 4. まとめ

参照点に依存した語意の学習手法を提案した。本実験では位置を表す単語や明るさを表す単語などを学習したが、今後は色を表す他の単語や、形を表す単語など、異なる属性を対象とする単語を同時に学習する実験を行う。

### 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] Iwahashi, N., Trans. IEICE, Vol. E91-D, No. 2, pp. 312-321.(2008)
- [3] Dempster, A.P., et al., Journal of the Royal Statistical Society. Series B (Methodological) 39 (1): 1-38.(1977)
- [4] Sugiura, K., et al., Advanced Robotics, Vol.25, No.6-7, pp.825-848.(2011)