

# ロボットによる動画像から抽出した 深層格情報を用いた言語獲得

植田 紗也佳<sup>†</sup> 岩橋 直人<sup>††</sup> 國島 文生<sup>††</sup>

<sup>†</sup> 岡山県立大学大学院情報系工学研究科

<sup>††</sup> 岡山県立大学情報工学部

## 1. はじめに

ロボットが人と対話するためには、実世界の事物と言葉との対応関係を理解している必要がある。人名・地名など、開発者が事前に想定できない未知のものに対応するため、ロボットにあらかじめ知識を与えるのではなく、周囲の環境に応じて自ら未知の事象の対応関係を学習することが望ましい。従来手法[1]では、名詞・形容詞・動詞のみから構成される構文で発話を行うことでロボットに学習をさせることが可能である。しかし助詞を含む自然言語を理解させるに至っていない。本研究では岩橋の研究[1]をもとに、意味役割(深層格)を示す格助詞を含む自然言語の意味をロボットに学習させることを目的とする。

## 2. 文と動画像内の深層格情報の結びつけ

発話を音節列に書き起こし、教師なし単語分割[2]により単語分割したテキストと動画像内の情報を深層格情報で表現したもののペアを1つの入力データとする。テキストと深層格情報の例を表1に示す。

表1 入力データの例

テキスト	こつぶ お はこ に のせて
深層格情報	Obj01_対象 Obj02_目標 NOSERU

この深層格情報はObj01という物体が対象格の役割、Obj02という物体が目標格の役割でNOSERUという動作を行うことを示している。なお深層格情報の順番は不同である。

### 2.1. 相互情報量を用いた対応関係の候補の取得

単語と物体・動作との相互情報量により、双方の対応関係を得る。各単語に対しすべての深層格情報との、相互情報量を求め、相互情報量の高い深層格情報を算出する。

## 3. 構文テンプレート生成

学習した単語と深層格情報のペアを元に、データ間で共通している構文パターンを見つけ出し、単語と深層格情報を記号に置き換えることで構文をテンプレート化する。相互情報量の高いペアから順に置き換えを行い、記号にどのような単語-深層格情報のペアが当てはまるかを記録する。これをすべてのデータについて行い、2回以上出現したテンプレートを有用なものとして取得する。

表2 テンプレートの例1

#001# お	#002# に	のせて
#001#_対象	#002#_目標	NOSERU

表3 テンプレートの例2

#001# に	#002# お	のせて
#002#_対象	#001#_目標	NOSERU

表2、表3の例では物体を記号に置き換えてテンプレート化している。例1と例2は同じ「のせる」という動作についてのテンプレートになるが、「お」「に」の助詞によって物体の役割が決められる。

## 4. 検証結果

### 4.1 対応関係の学習

単語20種類、物体10種類、動作6種類、言い回し3種類のデータを100個作成し学習させたところ、名詞・動詞には正しい物体・動作情報が割り当てられ、格助詞には正しい深層格情報が割り当てられた。

### 4.2 行動生成：文章から深層格情報を出力

学習結果を用いて50の文章を入力してそれぞれ対応する深層格情報を出力する実験を行ったところ、すべての文章に対して正しい深層格情報が出力された。

### 4.3 文生成：深層格情報から文章を出力

学習結果を用いて50の深層格情報を入力してそれぞれ対応する文章を出力する実験においても、すべての深層格情報に対して正しい文章が得られた。

## 5. 今後の課題

現在、データはすべて人手で作成したものである。よって今後は、発話の音声データを入力した時の音声認識誤りの問題と[3]の手法で動画像から得る物体・動作の特徴量情報を得た時のクラスタリングの問題に対処し、動画像から深層格情報を抽出するアルゴリズムを実装することで、実際の発話と動画像データからの言語獲得を実現する。

## 参考文献

- [1] 岩橋, 人工知能学会誌 18(1), pp. 49-58, 2003
- [2] 植田ほか, SSI2014, 岡山大学, SS3-3, Nov 2014
- [3] N. Iwahashi, IEICE TRANS. INF. & SYST., VOL. E91-D, No. 2, pp. 312-321, 2008