

# 環境情報を用いたロボットの言語理解

高瀬 健太<sup>†</sup> 石田 卓也<sup>†</sup> 秋山 俊介<sup>†</sup> 國島 丈生<sup>†</sup>

中村 友昭<sup>††</sup> 長井 隆行<sup>††</sup> 船越 孝太郎<sup>†††</sup> 中野 幹生<sup>†††</sup> 岩橋 直人<sup>†</sup>

<sup>†</sup> 岡山県立大 <sup>††</sup> 電通大 <sup>†††</sup> ホンダ・リサーチ・インスティテュート・ジャパン

## 1. はじめに

近年、家庭用ロボットの実用化に向けて広く研究が行われている。本稿では、家庭用ロボットの周囲の環境情報を利用して言語理解の精度を向上させる手法を提案する。

### 2.1. 提案手法

提案手法の概要を図1に示す。まず、入力された命令Xは音声認識され、認識結果 $Y_i$ に対して形態素解析を行う。次に認識結果を Bag of Words(BoW)表現に変換してそれを特徴ベクトルとして SVM によって行動クラス $a_n$ を識別する。また、CRF によりスロット抽出を行う。スロット抽出では、 $Y_i$ に含まれる名詞単語に IOB タグとその単語の属性 (ITEM, LOCATION, PERSON) をラベリングする。音声認識、行動識別、スロット抽出では、それぞれの n-best の認識結果とそのスコアを用いて識別結果の統合を行い行動とスロットを決定して行動生成を行っている。

提案手法では、CRF によるスロット抽出の結果に対して環境情報を利用する。環境情報とは、ロボットの周囲に存在する物や人、部屋の名前のことで、ロボットと人間の共通の知識として持つ実世界の情報のことである。スロット抽出によって、ラベリングされた名詞単語について、その単語が指し示すものが環境情報としてロボットが知っているものかを判断し、知らないものが含まれている命令は実行不可の命令として扱い、識別誤りとして棄却する。

### 2.2. 識別結果の統合

音声認識結果 $Y_i$ とそのスコア $S_{speech}(Y_i)$ 、行動識別結果 $a_n$ とそのスコア $S_{svm}(a_n)$ 、スロット抽出結果 $Z_j$ とそのスコア $S_{crf}(Z_j)$ の各 n-best を統合する。ここで、以下のようなスロットのスコア $S_{slot}(Z_j, a_n)$ を導入する。

$$S_{slot}(Z_j, a_n) = \frac{2PR}{P+R} + \phi$$

$$P = \frac{Z_j \text{のうち行動} a_n \text{に必要なスロット数}}{\text{行動} a_n \text{に必要なスロット数}}$$

$$R = \frac{Z_j \text{のうち行動} a_n \text{に必要なスロット数}}{\text{抽出されたスロット数}}$$

そして、これらのスコアから最終スコア S を最大に

する  $Y_i, a_n, Z_j$  を言語理解の結果とする。

$$S = S_{crf}(Z_j)^\alpha S_{svm}(a_n)^\beta S_{slot}(Z_j, a_n)^\gamma S_{speech}(Y_i)^\delta$$

ただし、 $\alpha=3.5, \beta=1.9, \gamma=7.9, \delta=6.6$  とした。

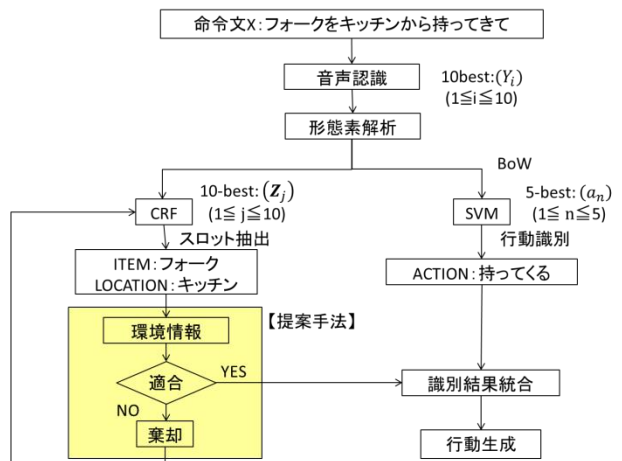


図1. 提案手法の概要

## 3. 評価実験

テストデータとして 600 文を用意した。従来手法(1)として、最終スコアに SVM, CRF, スロットのスコアを用いる場合、従来手法(2)として、それらに音声認識のスコアを追加した場合の 2 つを提案手法と比較する。図2に各手法の性能を示す。

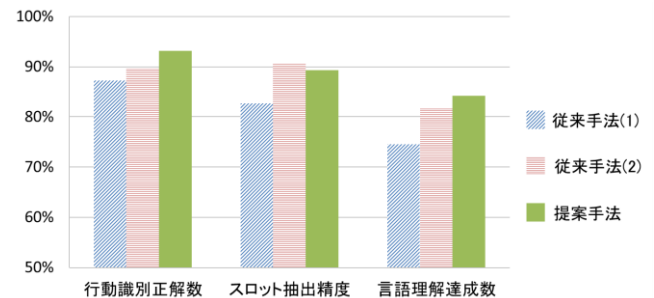


図2. 各手法の性能

## 5. 結論

評価実験より、家庭用ロボットの周囲の環境情報を利用することで言語理解の最終出力の候補を絞り、誤認識された音声認識結果や系列ラベリングにおけるスロット抽出の誤りを検出できた。これにより言語理解の精度が向上することが示された。

### 参考文献

[1] 中村友昭ほか, "SVM と CRF に基づくロボットによる自然言語理解" 第32回日本ロボット学会学術講演会, 2014