

音声を利用したロボットとのコミュニケーションに関する基礎研究

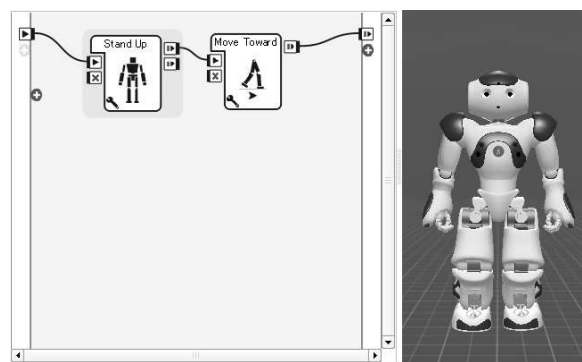
北見工業大学 ○桑岡 諒祐 鈴木育男 岩館健司 渡部美知子

要旨

近年の利用目的が多様化しているロボットに於いて、人間との直接的な関わりが増えてきている。そのため、操作者である人間とロボットの間で円滑なコミュニケーションを行う手段として音声認識の重要性が高まっている。そこで、その基礎的な研究として ALDEBARAN 社のヒューマノイドロボット NAO を使用し、簡単な英単語でのロボット制御を行った。

1. 研究の目的・意義

従来、ロボットの使用用途は人間の仕事を補佐するような産業用ロボットが主体となって発達してきた。しかし、近年では介護ロボットや掃除ロボットなど人間の生活に直接関わってくるロボットが増加しつつある。人間がロボットに命令を行い、ロボットが速やかに応答するといった人間とロボット間での円滑なコミュニケーションを行うには、そのためのシステムが必要となってくる。そこで、本稿では音声情報を利用したロボットの制御手法について検討する。



(a)プログラミングを行う画面 (b)シミュレーションを行う画面

Fig.1Choregraphe の画面の一部

2. 研究手法

本稿では、ALDEBARAN 社のヒューマノイドロボット NAO を使用する。また、システムを構築するソフトウェアに於いても ALDEBARAN 社により提供されているソフトウェア「Choregraphe」を使用する。

Fig.1 は、今回使用するソフトウェア「Choregraphe」の画像である。Fig.1(a)はプログラミングを行う画面であり、予め用意された「立つ」、「歩く」等の関数を配置することや、その関数を修正することで多彩なプログラミングを行うことができる。Fig.2(b)はシミュレーションを行う画面であり、Fig.1(a)の画面で作成したプログラムを実物のロボットを使わずに動作を確認することができる。

Fig.2 に、発話された言葉の認識から動作までのフローを示す。音声認識、単語の判別、各動作について 2. 1 節以降に示す。

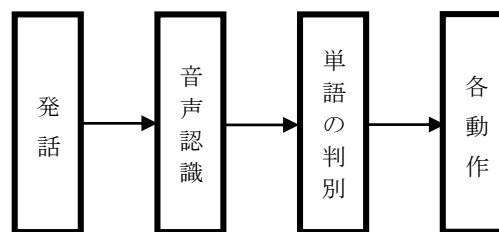


Fig.2 動作の流れ

2. 1 音声認識

「Choregraphe」では、始めに読み込ませる言語を設定し、NAO に搭載されているマイクを使用して音声情報の読み取りを行う。今回は、予め搭載されている言語の中で英語を使用する。また、予め NAO に認識できる単語として次の 5 語(yes,no,up,down,quit)を設定し、読み取った音声情報がこの内のどの単語に当たるかを判別させる。

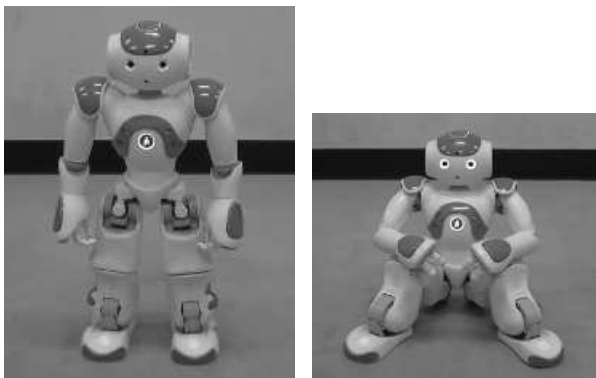
2. 2 判別と NAO の動作

音声認識によって得た文字列を使って、条件分岐により NAO の制御を行う。各単語と動作の関係について Table1 に示す。

Table1 各単語に対応する動作

単語	動作
yes	何も動作を行わない
no	何も動作を行わない
up	その場で立ち上がる
down	その場で座る
quit	プログラムを終了する

また、NAO が up と down の動作を行った時の画像を Fig.3 に示す。



(a)up 動作時の立った NAO (b)down 動作時の座った NAO

Fig.3 up, down 各動作時の画像

3. 実験手法

2章で提案した手法が正しく行われるかを確かめるため、次の実験を行った。設定した単語を話しかけ、正しく動作した回数を測定する。条件は以下の通りである。

- i. 1 単語につき各 10 回
- ii. 対象 5 人 (母国語を日本語とする人)

4. 実験結果

動作結果を Fig.4 に示す。

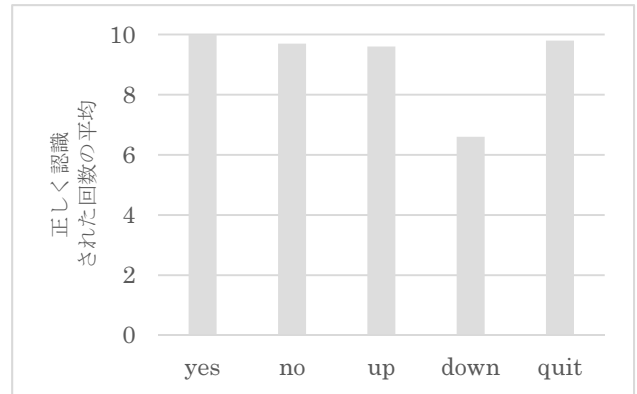


Fig.4 テスト結果

結果として yes, no, up, quit に関してはほぼ 100% 近く正しい動作を行うことができた。しかし、down に関しては、60% 程度しか正しく認識することはできなかった。また、各 5 人共 down をご認識する回数に大きなばらつきはなく、5 人共 50~70% 程度の正確さであった。また、down が誤って認識される場合全て no と認識されていた。この結果より、認識されやすい言葉、され難い言葉があることがわかった。

4. まとめ、今後の課題

今回は音声情報として、5 語の英単語を使用しロボットの制御を行った。その結果として、確実に認識することが困難な単語が見られた。今後日本語の文章での制御を目指す際も、同じように正しく認識され難い単語があることが予想される。そこで、日本語における認識され難い言葉への対応を考えていく必要がある。

参考文献

- [1] 田村優樹, 長崎達也, 中野雅広, 原田実, "意味解析に基づくロボット指示システム Athena2011", 情報処理学会研究報告. SLP, 音声言語情報処理 2012-SLP,91(10), pp.1-8, 2012
- [2] 戸塚典子, 伊藤彰則, "音声操作ロボットの動作とユーザ発話の音響的特徴との関係の分析", 電子情報通信学会技術研究報告. SP, 音声,113(161), pp.57-62, 2013