

音声命令言語と顔認識を用いた復興作業支援ロボット -命令認識システムの開発と評価-

†内野 翔 ‡岡 哲資

†日本大学大学院生産工学研究科数理情報工学専攻

‡日本大学生産工学部数理情報工学科

1. はじめに

近年、世界中で地震や洪水などの自然災害が多発している。特に日本は地震大国と言われており、地震の被害は多く見られる。このような自然災害後の復興作業は人による作業が主であり、人員不足による作業の遅れが目立っている。

そこで今、ロボットを利用した復興作業に目が向けられている。現在、活躍しているロボットの主流は、遠隔操縦型の探査ロボットであり、地上走行型ロボット[1]、飛行型ロボット、水中移動型ロボットである[2]。その他、大きなものを運ぶための遠隔操作型ロボット[3]が挙げられる。

しかし、これらのロボットはコントローラーを用いた複雑な操作が要求されるため、使用できる人員が少ない。また、長いものなど一台で持つのが困難な物は二台のロボットで運ぶ必要がある。二台のロボットを一人で操作することはより一層困難であり、二人で操作する方法では必要な人員の数は変わらない。この問題を解決するため、本研究では人が二人で行っていた物を運ぶ作業を一人とロボット一台の共同作業で実現することを目指している。

物を運ぶ共同作業では、ユーザは両手が塞がっているため、コントローラーが使用できない。したがって、本研究では、音声命令言語と顔認識を用いる。音声は、Siri、しゃべってコンシェル[4]などのスマートフォンでの操作で実用化されており、たけまるくん[5]、MMD Agent[6]などの音声対話の研究も進んでいる。語彙と文法規則により入力を限定した命令言語によって、高い認識率が実現できると考えられる。

Human-robot cooperative work using spoken command language and face tracking in post-disaster recovery -development and evaluation of command recognition system-

†Sho Uchino,

Mathematical Information Engineering, Graduate School of Industrial Technology, Nihon University

‡Tetsushi Oka,

Mathematical Information Engineering, College of Industrial Technology, Nihon University

また、顔認識を併用することで、雑音による音声の誤検出によってロボットが誤動作することを防ぐ。

本稿では、共同作業を復興支援作業ロボットに必要な、音声命令言語と顔認識を用いた命令認識システムの開発と評価について報告する。

2. 命令認識システム

前述の通り、人との共同作業によって物を運ぶロボットのための命令認識システムを開発した。本システムは、音声認識と顔認識を組み合わせることにより、ユーザがロボットの方を向いて発した音声のみを命令と解釈する。

本システムが認識する音声入力は、「前進」、「後退」、「右に回れ」、「左に回れ」、「右移動」、「左移動」、「止まれ」の7種類である。これらは、全てロボット視点での命令になっている。例えば、ロボットを前に動かしたい場合は「前進」と言う。「右に回れ」と「左に回れ」はロボットを中心にユーザが弧を描く移動をしたい場合に使用する。

ユーザがロボットの方を向いているかどうかは、ユーザの頭の追跡と正面の顔の検出を用いて判定する。顔検出は、画像の頭の座標近くの領域のみで行う。

命令認識システムのハードウェアは、Logicool社の無線ヘッドセットマイクH800、ノートパソコン一台(Windows7 32bit, メモリ 2GB, Intel Core i5搭載)、RGB-DカメラであるKinect for Windowsを使用して実現した。

ソフトウェアは、音声認識に大語彙連続音声認識エンジンJulius4.3.1[7]、顔認識とKinect画像出力にOpenCV2.3.1、VisualStudio2010を使い、C++で開発した。

3. 命令認識システムの評価方法

8人のユーザによる命令認識システムの評価を行った。各ユーザにつき、次の順序で実施した。

1. ユーザにシステムの仕組みを説明する。
2. Kinectから約1.5[m]離れて立ってもらう(図1)。
3. Kinectとは違う方向を向いてもらう。
4. 前述の7つの命令を2回ずつ、正しく認識されるまで行ってもらう。



図1. システムとユーザの位置関係

4. 評価結果

命令総数は139回で、そのうちシステムが正しく認識できたのは112回(認識率81%)であった。命令別では、「右移動」が51%、「前進」が72%、「左に回れ」が78%、「右に回れ」が88%、他は全て100%であった。認識失敗は、音声の聞き逃し(21回)と誤認識(6回)によるものであり、命令時の顔認識の見落としはなかった(図2)。6回の誤認識はいずれもシステムが「止まれ」と認識したものである。

音声の誤検出は1回あり、「止まれ」で認識した。評価中、正面の顔が検出されていないとき、音声入力が26回検出されていた。全ての命令を失敗なく行えたユーザは4名であった。「右移動」を100%認識できたユーザは5人であった。

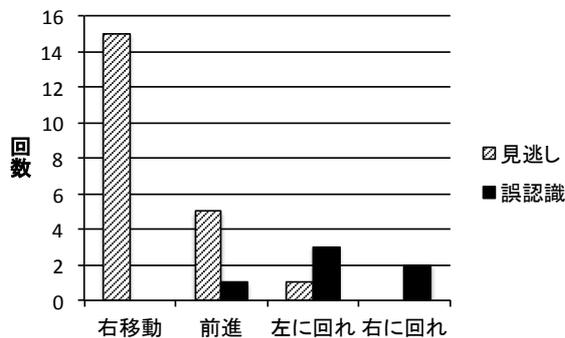


図2. 認識失敗の内訳

5. 考察

誤検出数と認識失敗の内訳をみると、本システムによって、ロボットが誤動作する可能性は低いといえる。また、「止まれ」の認識率は100%であり、動作中のロボットを確実に止めること

ができる。しかし、認識率の低い命令については、何度も命令しなければロボットが動作しないため、作業の効率上問題である。聞き逃しが起きやすい原因を解明し、システムを改良する必要がある。また、ユーザによって認識率の違いがあることから、ユーザごとに認識しやすい音声命令言語を設定することも解決策として考えられる。

6. おわりに

本稿では、復興作業ロボットとの共同作業のための命令認識システムの開発と評価について述べた。評価結果から、開発したシステムを用いることで共同作業におけるロボットの誤動作が起きにくくなることが示された。しかし、命令ごとの認識率からはまだ十分効率的な共同作業が行えるとは言えず、音声命令言語とシステムの改良が必要である。

今後、音声命令言語の改良・拡張を行い、命令認識システムを組み込んだロボットの開発と評価を行う予定である。

参考文献

- [1] 小柳 栄次：サブローラを持つレスキューロボット，日本ロボット学会誌，Vol. 28, No. 2, pp. 147-150, 2010
- [2] 大野 和則：レスキューロボットにおけるテレロボティクス，日本ロボット学会誌，vol. 30, No6, pp. 571-573, 2010
- [3] ロボナブル，無人化施行技術・遠隔管理を駆使して作業
<http://www.roborable.jp/trend/2011/04/16cat.html>, 2015
- [4] 西村 竜一，原直，川波 弘道，李 晃伸，& 鹿野 清宏：10年間の長期運用を支えた音声情報案内システム「たけまるくん」の技術(〈特集〉音声対話システムの実用化に向けて)，人工知能学会誌，Vol. 28, No1, pp. 52-59, 2013
- [5] 吉村 健：しゃべってコンサルと言語処理，情報処理学会研究報告．SLP，音声言語情報処理，No. 4, pp. 1-6, 2012
- [6] 李 晃伸，大浦 圭一郎，& 徳田 恵一：魅力ある音声インタラクションシステムを構築するためのオープンソースツールキット MMDAgent オーガナイズドセッション：音声対話，第13回音声言語シンポジウム，電子情報通信学会技術研究報告SP，音声 111, pp. 159-164, 2011
- [7] 河原 達也，& 李 晃伸：連続音声認識ソフトウェア Julius，人工知能学会誌，Vol. 20, No. 1, pp. 41-49, 2005