

# 音声と手のジェスチャーによる 家庭用ロボットへの家電操作命令

島津 正次郎<sup>†</sup> 岡 哲資<sup>‡</sup> 中村 喜宏<sup>‡</sup>

日本大学大学院生産工学研究科数理情報工学専攻<sup>†</sup>

日本大学生産工学部数理情報工学科<sup>‡</sup>

## 1. はじめに

近年、家庭への無線ネットワークの普及によって、同じ場所から様々な情報家電を操作できるようになりつつある。家庭を智能化するスマートホームの研究は、以前から行われており[1]、スマートフォンからの情報家電の操作も実現している[2]。また、音声認識技術などの進歩により、音声を利用したスマートフォン上のアプリケーション・インタフェース(Apple の Siri や NTT ドコモの「しゃべってコンシェル」[3]など)も注目されている。しかし、多機能なタブレットやスマートフォンを使いこなせない人、操作が複雑であるというイメージなどから使いたがらない人もいる。また、家電を操作したいときにスマートフォンやリモコンが見つからない状況も考えられる。

一方、お掃除ロボットルンバ[4]などの家庭用ロボットも普及し始めている。家庭用ロボットは、今後掃除以外にもさまざまな目的で利用されるものと予想される。家庭内にロボットがいれば、スマートフォンやリモコンが手元になくても、情報家電の操作をロボットに命令すればよい。また、ロボットを呼び寄せることができれば、移動せずに命令が行える。さらに、魅力的で親しみやすいロボットやキャラクターが開発されれば、スマートフォンや GUI になじまない人も積極的に家電操作に利用する可能性がある。

著者らは、家庭用ロボットを音声とジェスチャーを組み合わせる命令方法を研究している。例えば、手のジェスチャーで位置や移動量を伝え、音声と組み合わせる呼び寄せ、適切な距離を保つ方法についての研究例がある[5]。

音声とジェスチャーを用いれば、手元の端末が不要になる。また、多くの場合、GUI 操作に比べて容易に命令ができると考えている。音声にジェスチャーを加える利点は、雑音などによる音声の誤検出の除外、長さ・方向や三次元位置などの音声では伝えにくい情報の伝達、冗長性による認識率の向上などが考えられる。本発表では、家庭用ロボットと向き合った状態での家電操作の具体的な命令方法、命令認識システムの実現方法、命令の認識率及び誤反応の起きやすさの評価の方法と結果を示す。

## 2. 命令方法

ロボットへの家電操作の命令は、右手または左手を顔の高さに上げて止めるジェスチャーと「テレビをつけろ」などの音声を組み合わせたものである。発話と手のジェスチャーを同時に行って命令することで、ひとりごと、他の人への発話、雑音による音声の誤検出を除外する。

表 1 に各命令の音声及びジェスチャーを示す。「つけろ」、「上げろ」などの音声に右手、「消せ」、「下げろ」などの音声に左手のジェスチャーを割り当てた。「つけろ」・「消せ」など、対になった命令を左右の手に割り当て、正を右手、負を左手とすること、体を動かしながら命令することで、分かりやすく、覚えやすくすることがねらいである。また、「上げろ」と「下げろ」の聞き間違いや不明瞭な音声にも対応できる。「かえろ」に対しては、左右どちらの手を用いてもよい。

同じ家電の操作を続けて命令する場合は、「テレビの」などの家電の指定は省略してもよい。チャンネル変更、FM/AM 切り替えにおいても家電の指定は省略できる。温度、風量、チャンネルを変更する命令では、手を上げている間、設定値が変化し、手を下したタイミングで値を確定する。設定値は家庭用ロボットの表示画面に示され、ロボットのユーザは、これを確認しながら手を下す。

---

Commanding a domestic robot by speech and hand gesture to operate home information appliances

<sup>†</sup>Syojiro Shimadzu,

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

<sup>‡</sup>Tetsushi Oka, Yoshihiro Nakamura,

Mathematical Information Engineering, College of Industrial Technology, Nihon University

表 1 家電操作の命令方法

音声	手
テレビ(ラジオ/冷房/暖房/エアコン)をつける	右
テレビ(ラジオ/冷房/暖房/エアコン)を消せ	左
テレビ(ラジオ)の音量を上げる	右
テレビ(ラジオ)の音量を下げる	左
冷房(暖房/エアコン)の温度を上げる	右
冷房(暖房/エアコン)の温度を下げる	左
冷房(暖房/エアコン)の風量を強くしろ(上げる)	右
冷房(暖房/エアコン)の風量を弱くしろ(下げる)	左
テレビのチャンネルを変える	右(左)
ラジオをFM(AM)に変える	右(左)

### 3. 命令認識システム

前節で述べた音声とジェスチャーを組み合わせた命令を認識するシステムを開発した。音声認識には、大語彙連続音声認識エンジンJulius 4.2 (<http://julius.sourceforge.jp/>) を用いた。Juliusの認識文法の記述によって、表1の命令の音声を家電の省略も含めて全て認識できるようにした。ジェスチャー認識には、視覚センサMicrosoft KinectとSDKのSimpleOpenNI 0.27 (<http://code.google.com/p/simple-openni/>) を用いた。SimpleOpenNIのユーザ発見と骨格トラッキング機能を利用し、頭、肩、肘、手先の三次元位置を追跡し、手のジェスチャーの認識を実現した。本システムは、両手を高く上げたユーザに注目し、注目しているユーザのジェスチャーのみを検出する。また、命令の音声とジェスチャーが近い時刻に検出されたときのみ命令と認識するため、音声及びジェスチャーの誤検出があったとしても、同時に発生しない限り無視される。

### 4. 評価実験

#### 4.1 実験方法

開発したシステムにおける命令の認識率及び誤反応の起きやすさを複数のユーザで評価する。まず、音声と手のジェスチャーを同時に組み合わせて命令することを伝える。次に、図1のロボットと向き合った状態で表1の命令を行ってもらう。ロボットの距離は1.2-2[m]程度とする。ロボットの液晶ディスプレイに、各命令の言葉と左右どちらの手を上げるかを表示し、その通りに命令してもらう。また、手を下すタイミングも合図する。音声とジェスチャーの認識結果と時刻を記録し、ユーザが命令している様子をビデオで撮影する。

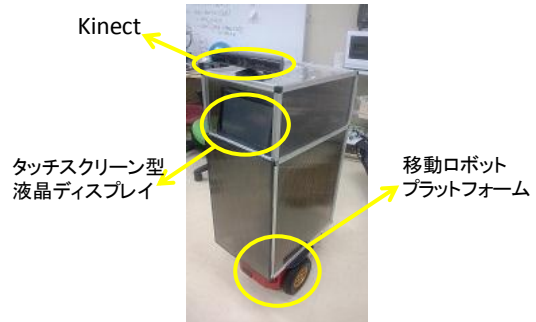


図 1 家庭用ロボット RockyII

#### 4.2 実験結果

8人に評価実験を行った。ジェスチャーに関しては6人が100%認識した。1人は、命令する間で小さく動き誤検出された。もう1人は命令を言い間違えたとき手を上げたため誤検出された。音声に関しては4人が100%認識した。他の4人は2~6個誤認識された。誤認識された音声は「上げる」が「下げる」、「下げる」が「上げる」といったことがあった。また、音声として認識されないことがあった。

### 5. 考察

ジェスチャーは8人中6人100%認識したことから良いと考えられるが小さな動きにも反応するため、反応する条件をもっと厳しくする必要がある。音声に関しては誤認識が多かったため、命令の音声を言いやすい言葉に変更するなど工夫が必要である。

### 6. まとめ

音声と手のジェスチャーによる移動ロボットへの家電操作命令方法、命令認識システムの実現方法、評価実験の方法と結果について述べた。実験結果からジェスチャーの認識は良かったが、音声の認識は誤認識が多くみられた。今後の課題には、命令の容易さの評価、覚えやすさ・忘れにくさの評価、他の命令方法に対する優位性の検証などが挙げられる。

謝辞 本研究は科学研究費補助金(22500179)からの補助を受けた。

#### 参考文献

[1] Li Jiang: "Smart Home Research" Proc. of 2004 International Conf. on Machine Learning and Cybernetics, pp.659-663 (2004)  
 [2] パナソニック スマート家電 <http://panasonic.jp/pss/ap/>  
 [3] 辻野孝輔 栄藤稔 磯田佳徳 飯塚真也「実サービスにおける音声認識と自然言語インタフェース技術」人工知能学会誌 28-1 pp.75-81 (2013)  
 [4] iRobot社 ルンバ <https://www.irobot-jp.com/roomba/>  
 [5] 木林龍一「音声と手の三次元動作を用いた生活支援ロボットとのマルチモーダル対話」日本大学生産工学研究科数理情報工学専攻修士論文 (2012)