

平成 21 年 6 月 6 日現在

研究種目：基盤研究 (C)  
 研究期間：2007～2008  
 課題番号：19500171  
 研究課題名 (和文) 家庭用ロボットのマルチモーダル命令言語の開発  
 研究課題名 (英文) Development of a multimodal command language for home-use robot users  
 研究代表者  
 岡 哲資 (OKA TETSUSHI)  
 福岡工業大学・情報工学部・准教授  
 研究者番号：00282921

## 研究成果の概要：

特別な訓練を受けずに短期間で習得できる家庭用ロボットの命令方法として、手振り、ロボットの体の接触、ボタン操作と日本語音声を組み合わせたマルチモーダル（複数の伝達様式を用いる）命令言語を開発し、仕様の詳細を明らかにした。その有効性と問題点を調べるため、留守宅確認、情報家電操作、掃除、箱の移動などを命令できる家庭用ロボットを実現した。100人以上による評価試験を行い、90分以内で命令言語の習得が可能であることを確認した。

## 交付額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	1,000,000	300,000	1,300,000
2008年度	900,000	270,000	1,170,000
年度			
年度			
年度			
総計	1,900,000	570,000	2,470,000

## 研究分野：総合領域

科研費の分科・細目：情報学 知覚情報処理・知能ロボティクス

キーワード：物理エージェント、家庭用ロボット、マルチモーダル、命令言語

## 1. 研究開始当初の背景

幅広いユーザを対象とした家庭用ロボットの命令を可能にするには、パーソナルコンピュータやビデオレコーダ、携帯電話などで用いられるグラフィカルユーザインタフェース(GUI)、二次元バーコード認識、音声認識と自然言語理解などの技術が利用可能であるが、ユーザにとっての容易さ、実現コストと即時性の問題を同時に解決した実用化は難しかった。GUIや二次元バーコードでの命令は比較的低コストで実現可能であるが、家庭用ロボットへの多種多様な命令を可能

にするには、習得に時間がかかり、操作の手順が複雑でユーザへの負担が大きく、即時性に欠ける。一方、日本語などの自然言語の音声による命令は、新たな技能の習得が不要であり、ユーザの負担を軽減できるが、自然言語理解のための計算コストが高く処理時間が長いこと、自然言語の発話が曖昧であることが問題である。

人間同士のコミュニケーションにおいては、言語に加えて、声の調子、手振り、視線など様々な伝達様式が用いられており、言語のみでは情報伝達が難しい場合もある。以上

から、家庭用ロボットの実用化のためには新しい命令方法の確立が必要であった。

## 2. 研究の目的

本課題の目的は、ユーザの負担と実現コストおよび即時性の問題を同時に解決し、かつ自然言語の持つ曖昧さなどの弱点を排除した家庭用ロボットの命令言語の一例を示し、その有効性と問題点を明らかにすることである。命令言語は、日本語話者を対象とし、音声言語とボタン操作、ロボットの体の接触、手のジェスチャを組み合わせたマルチモーダルな（二つ以上の伝達様式を用いる）ものである。

## 3. 研究の方法

特別な訓練を受けずに短期間に習得できる家庭用ロボットの命令方法として、片手・両手の手振り、ロボットの体の接触、ボタン操作と日本語の音声を組み合わせた命令言語を開発した。次に、留守宅確認、家電操作、掃除、物体移動などを本言語で命令できる家庭用ロボットを実現した。予備知識のないユーザ100人以上によるロボットの利用試験によって、命令言語の有効性を調査した。

まず、家庭用ロボットへの命令の内容を吟味し、ユーザに覚えやすく、計算コストが低い、曖昧さのない音声命令言語の文法を語彙(250語以上)と300以上の文法規則の形で作成し(表1)、音声認識ソフトウェアによって命令の認識を可能にした。可能にした音声命令は、前進・後退、旋廻、サイドステップなどの個別の移動、目的地までの移動、視線方向の変更、情報家電の操作、掃除、物体の把持と設置など、家庭用ロボットに必要な行為20種類以上に、複数のパラメータ属性(速さ、角度など)を加えたものである(表2)。例えば、「大きく右に旋廻」、「ゆっくり45度回れ」、「少し前進」、「エアコンの温度を28度に設定しろ」、「冷蔵庫に移動しろ」、「床から幅30センチの箱を持ち上げる」などのように、属性を数値的または定性的に指定できるようにした。命令の曖昧さを排除するため、文法規則外の命令は受け付けず、属性を省略した場合、ユーザの意図を推論するのではなく、デフォルト値を用いて解釈を行うこととした。

表1 命令言語の構文規則の例

規則	説明
S → ACTION	行為命令
S → MODIFIER	命令の修正
ACTION → AC2	クラス2の行為命令
AC2 → P2 AT2	属性と行為の型
P2 → SPEED DIST	属性(速さ、移動量)

P2 → SPEED	属性(速さのみ)
DIST → NUMBER LUNIT	数値、長さの単位
SPEED → SPEEDW	速さを表す語

音声命令文法の開発と合わせて、表2のような行為の種類と属性を含む命令の計算機上の意味表現形式を決定し、音声認識結果(発話中の単語の列)の意味解析を行うプログラム(命令解釈モジュール)を開発した。

表2 命令の種類とパラメータ属性の例

行為の種類	属性
前進・後退	移動量・速さ
旋廻	方向・速さ・角度
サイドステップ	方向・速さ・移動量
視線変更	方向・速さ
エアコン温度設定	温度(摂氏)
エアコン温度変更	温度差(摂氏)
物体把持	高さ・幅
物体設置	高さ
掃除	広さ・丁寧さ・モード
目的地への移動	目的地

次に、この音声命令言語をマルチモーダル命令言語拡張した。音声言語の命令の一部を片手および両手の往復ジェスチャ(速度、振動数、振幅、往復回数、両手の距離)、ロボットの肩、腕、手先への接触時間、ボタンの押下(種類・押し時間・回数)で代用する、マルチモーダルな命令方法を考案した。例えば、片手を右方向に高い振動数で三往復させ、「回れ」という発話を行うと、「右に速く90度回れ」と同じ意味になるように、ノンバーバルな(言葉によらない)メッセージを行為の属性パラメータと対応付けた。表3は、手のジェスチャと属性の対応付けである。また、図1に12個のボタンの配列とそれに対応する属性を示す。表4のように、ボタン操作の押し時間と回数も属性と対応付けた。音声で省略された属性以外に、「ここ」、「こっち」、「このくらい」などの語と合わせてジェスチャや接触を利用できるようにした。

表3 片手および両手ジェスチャの利用

行為	属性	ジェスチャ特徴量
旋廻	方向	方向
	速度	速度・振動数
	角度	振幅・回数
前進・後退	移動量	振幅
	速度	速度・振動数
把持	高さ	往復回数
	幅	両手の距離
設置	高さ	往復回数





手引きを参照しつつ、自発的にマルチモーダル命令を行い、与えられた目標を達成するものである。ユーザの様子をビデオ録画し、音声認識結果、実行された命令、タスク完了までの時間などを記録した。課題の実施後には、ユーザの印象、好み、命令のしやすさ、問題点などを調べるためのアンケートを取った。説明、練習、課題実施、アンケートを含め、全体を60分または90分以内で実施した。

#### 4. 研究成果

留守宅確認、家電操作、物体の移動、掃除などのユーザ評価を通じ、開発したマルチモーダル命令言語の正しい利用が、予備知識のないユーザによって、90分以内で習得可能であることが確認できた。ジェスチャの利用方法やボタンの押し時間など、練習が必要なものもあり、一部習得が難しい部分があったものの、命令の成功率は90%以上と高いものであった。汎用パーソナルコンピュータ上で実現した命令解釈システムが、マルチモーダル命令の意味解釈に要する時間は1秒以内であった。ユーザがジェスチャやボタン操作と発話に要する時間とシステムの情報処理に要する時間を含めた、個々の命令の所要時間は大半が5秒以内、長くても10秒以内に収まっていた。

以上の結果より、開発した命令言語が家庭用ロボットの実用化に有効であり、従来のGUIよりも習得の容易さと命令の所要時間の両面で優ることが示された。また、語彙と文法規則を制限し、曖昧性を排除したことによる大きな問題は発生せず、逆に高い音声認識率、低い計算コストを実現することができた。以上より、開発した命令言語の自然言語(日本語)に対する優位性が示された。

この他、ユーザ評価の記録を分析することで、家庭用ロボットの実用化に関連する様々な知見が得られた。主なものを以下で述べる。

まず、マルチモーダル命令を好むユーザが、音声のみの命令を好むユーザを大きく上回った。また、マルチモーダル命令は語数が少ないことから、音声認識誤りのリスクを減らすことも確認された。

一方で、ボタンやジェスチャで音声の一部分を代用する場合の対応付けが自然かつ適切なものでないと、命令方法の習得が難しくなることから、対応付けの自然さ、命令方法の説明や練習方法などの工夫の重要性が確認された。特に、移動量や角度を音声と同様に多段階(アナログ)で指定するには、さらなる工夫が必要である。また、手のジェスチャ入力をより容易にするには、手の移動量や

両手の距離を一定以上の精度で計測するための三次元視覚認識が重要であることが分かった。

音声言語のみの命令は、語数が多い命令には多少の慣れが必要であり、音声認識誤りのリスクがあるが、命令方法自体は短時間で習得できることが分かった。また数字を含む発話は、認識誤りがやや多くなった。音声認識率は、認識エンジンやマイクロフォンの性能に依存するが、語彙と文法の制限により、最新の技術で実用上十分高い認識精度が得られる見通しを得た。

この他、ユーザ評価結果から、課題の完了時間には、個々の命令の成功率だけでなく、適切な命令を選択する能力、命令の適切なタイミングの理解も必要になることがわかった。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計2件)

- ① 岡 哲資、阿倍豊和、杉田薫、横田将生、RUNA: a multimodal command language for home robot users, Artificial Life and Robotics, 13, 455-459, 2009, 査読有り
- ② 岡 哲資、阿倍豊和、杉田薫、横田将生、Success rates in a multimodal command language for home robot users, Artificial Life and Robotics, 14, 掲載決定, 2009年-2010年, 査読有り

[学会発表] (計7件)

- ① 阿部豊和、岡 哲資、杉田薫、横田将生、Success rates in a multi-modal command language for home robot users, 14<sup>th</sup> International Symposium on Artificial Life and Robotics (AROB14), 2009年2月5日、ビーコンプラザ(大分県別府市)
- ② 阿部豊和、岡 哲資、杉田薫、横田将生、Multi-modal interface with remote home robots, 14<sup>th</sup> International Symposium on Artificial Life and Robotics (AROB14), 2009年2月5日、ビーコンプラザ(大分県別府市)
- ③ 岡 哲資、阿倍豊和、下地昌太、中村拓



哉、杉田薫、横田将生、Directing humanoid in a multi-modal command language、17<sup>th</sup> IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN08)、2008年8月3日、ミュンヘン工科大学(ドイツ ミュンヘン)

- ④ 岡 哲資、阿倍豊和、下地昌太、横田将生、音声とボタンを組み合わせた命令言語によるヒューマノイドの遠隔操作、日本機械学会ロボティクス・メカトロニクス講演会 08、2008年6月5日、ビッグハット (長野県長野市)
- ⑤ 岡 哲資、阿倍豊和、中村拓哉、横田将生、マルチモーダル命令言語 RUNA による小型ヒューマノイドの操作、日本機械学会ロボティクス・メカトロニクス講演会 08、2008年6月5日、ビッグハット (長野県長野市)
- ⑥ 岡 哲資、阿倍豊和、杉田薫、横田将生、RUNA: a multimodal command language for home robot users、13<sup>th</sup> International Symposium on Artificial Life and Robotics (AROB13)、2008年1月31日、ビーコンプラザ(大分県別府市)
- ⑦ 岡 哲資、佐々木崇、簡素な音声言語と触れ合いで指示が可能なロボット、日本機械学会ロボティクス・メカトロニクス講演会 07、2007年5月10日、アルヴェエ (秋田県秋田市)

## 6. 研究組織

### (1) 研究代表者

岡 哲資 (OKA TETSUSHI)  
福岡工業大学・情報工学部・准教授  
研究者番号： 00282921

### (2) 研究分担者

なし

### (3) 連携研究者

なし