

タッチパネルと音声による生活支援ロボットの マルチモーダルインタフェース

日大生産工(学部) ○松本 大賢
日大生産工 岡 哲資

1. はじめに

近年、ロボットの研究が進んできたことで家庭用ロボットの普及が期待されている。しかし、家庭用ロボットはすでに普及している産業用ロボットに比べ、多くのことが求められる。特に、人の生活を手助けする生活支援ロボットは、幅広いユーザがロボットに思い通りのタスクを実行させられることが必要である。

本研究の目的は、生活支援ロボットに対して、ロボットやPCの知識を持たない人でも、直感的に命令ができるインタフェースを開発し、その有用性について検討することである。

現状では、ロボットに対するインタフェースは音声によるものが一般的と言われる。しかし、音声を正しく認識し、かつ適切な行動を生成することは困難であり、また空間中の位置など、音声では表現しにくい情報も存在する。そのため、音声以外によるインタフェースの研究が数多くされている。例えば、実世界へカードを置くことによる指示や実世界画像上での操作による指示、レーザジェスチャーによる指示などがある¹⁾。

また、音声とその他の入力を組み合わせたマルチモーダルインタフェースも研究されている。現状で実現しているマルチモーダルインタフェースは、音声とボタン入力によるもの、音声と手を使ったジェスチャーによるもの、音声とロボットのボディに触れることによるものなどがある²⁾³⁾。

従来のマルチモーダルインタフェースでは、環境中の任意の3次元位置を伝える方法がない、ロボットが命令を実行しているときには、命令の変更などの情報を伝えることが十分にできない、システムがユーザの命令をどのように解釈したか分からない、といった問題点がある。

本研究で実現するマルチモーダルインタフェースは、音声とタッチパネルによるものと

する。実現するマルチモーダルインタフェースでは、タッチパネルへのタッチにより、任意の三次元位置を伝えることができるように工夫する。また、ロボットが命令を実行しているときに、タッチパネルにボタンを表示し、追加の入力をユーザに促す。さらに、合成音声によってユーザに命令をどのように解釈したのか伝えるようにすることで、従来の問題点を解決する。

想定する使用方法には、ユーザから離れたところにある物を手元まで持ってこさせる、手元にあるものを指定した場所まで持って行かせる、空き缶やペットボトルをゴミ箱まで捨てに行かせる、が挙げられる。

2. インタフェース

2.1. タッチパネルと音声を組み合わせた命令

実現するインタフェースでは前進、後退、移動、旋回、物を掴む、物を置く、物の移動の7種類の命令ができる。ロボットシミュレーションソフトwebots上の仮想ロボットに対する命令を行うものとし、仮想ロボットは、ステレオカメラとGPSを積んだ、アーム付車輪移動ロボットとする（図1）。

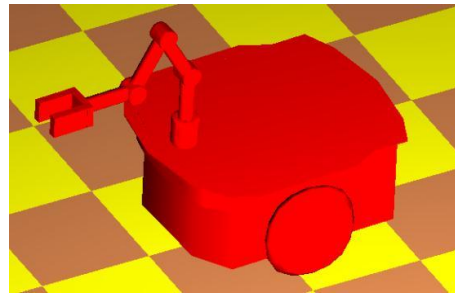


図1 使用する仮想ロボット

タッチパネルの画面には、ロボットのステレオカメラの右のカメラ画像と左のカメラ画像、ボタンがそれぞれ別のウィンドウに表示

Multimodal Interface with Life-supporting Robots
based on Speech and Touch Screen System

Hirosato MATSUMOTO, and Tetsushi OKA

される。表示されるカメラ画像をタッチすることで、ロボットを中心とした三次元位置をロボットに伝えることができるよう工夫する。また、ボタンは必要な時に必要なものだけを表示する。

2.2. 座標入力を用いた命令

座標入力とは、目標となる物体や場所の三次元位置をロボットに伝えること、とする。

実現するインタフェースでは、座標入力はカメラ画像をタッチすることで行う。

座標入力を用いると、例えば、目標位置の座標入力を行い、音声で「ここまで進め」と入力することで命令できる。また、目標とする方向の座標入力を行い、音声で「こっちを向け」と入力することで命令できる。さらに、目標となる物と、目標位置の座標入力を順番に行い、音声で「ここからここまで移動」と入力することで命令できる。以上の様に目標となる物体の位置座標入力を利用する。

2.3. 曖昧な入力への対応

音声による入力は、曖昧な入力となる場合がある。例えば、「前進」とだけ入力があったときは、どこまで前進するのか、「回れ」とだけ入力があったときは、左右どちらに旋回するのかが曖昧である。

曖昧な入力に対しては、ボタンを表示して追加の入力を促すことで、対応する。「前進」とだけ入力があったときは、停止ボタンを表示して、前進を始め、停止ボタンがタッチされた時点で停止する。「回れ」とだけ入力があったときは、右ボタンと左ボタンを表示し、ユーザからの次の入力があるまで待つ。

2.4. 命令の確認

音声認識の認識誤りやユーザの言い間違いは完全には避けられない。認識誤りや言い間違いが起こったときに、すぐに命令の回復ができるようユーザへの確認を行う。確認は、命令内容を合成音声でユーザに伝え、OKボタンとキャンセルボタンを表示し、命令が正しく伝わったか、ユーザの判断を求める。

2.5. 初心者向けの命令認識

初心者が音声入力を行うとき、所々で言いどみ、音声入力がとぎれとぎれとなり、認識が行えない場合があることが考えられる。

音声入力がとぎれとぎれであった場合も認識が行われるよう工夫する。

3. 実験システム

実験に使用するシステムにおいて、特に重要なタッチイベント検出と音声認識について説明する。

タッチイベントの検出では、一方のカメラ画像をタッチすると、もう一方のカメラ画像と照らし合わせて対応点を求め、三次元位置を算出するようにする。また、ボタンをタッチしたときと、指を離れたときにイベントを検出することで、ボタンを押している間、アームを動かすといったことができるようにする。

音声認識は、一単語ごとに認識できるようにし、とぎれとぎれの音声入力でも認識できるようにする。

4. 実験内容

実験では、一般家庭をモデルにした仮想環境を作成する。この仮想環境の中で、様々なユーザに、複数のタスクを課題として提示し、ロボットに命令してもらう。課題は、「見えているものを掴ませる」や「見たいものが見やすい位置へロボットを移動させる」といったものを提示する。その後、アンケートをとり、作成したログとアンケートを使って、実現するインタフェースの有用性を検討する。

5. おわりに

タッチパネルと音声による生活支援ロボットのマルチモーダルインタフェースを実現する。タッチパネルと音声を用いて、従来のインタフェースに比べ、より直感的な命令を行えるように工夫する。実験により有用性を検討する。

「参考文献」

- 1)五十嵐健夫, “解説 ロボットのためのユーザインタフェース”, 日本ロボット学会誌, Vol.28, No.3, (2008), pp.246-248
- 2)Tetsushi Oka, "Multimodal Command Language to Direct Home-use Robots", V.A.Kulyukin ed., Advances in Human-Robot Interaction, INTECH, (2009), pp.221-232
- 3)Tetsushi Oka, Kaoru Sugita and Masao Yokota, "Commanding a humanoid to move objects in a multimodal language", Artificial Life and Robotics, Vol.15, No.1, (2010), pp.17-20