

音声と直指操作による
ロボット制御のシミュレーションシステム

1S-5

張 建新 中川 聖一
豊橋技術科学大学

1. はじめに

近年、音声応用インタフェースのひとつとして、複数の情報入力手段を利用したユーザインタフェースであるマルチモーダル入力インタフェースの研究が進められている[1,2,3]。本稿で、音声と直指操作による2つの入力手段を用いたロボット制御のシミュレーションシステムを作成し、このシステムに対して、実験者の反応や入力時間の評価および入力が曖昧な場合において、従来の入力方法（音声だけあるいは直指操作だけ）との比較実験を行なった。最後に、音声と直指操作を入力手段としたマルチモーダル入力システムの利点および問題点について考察した。

2. 音声認識

2.1 認識方法

- ◎サンプリング周波数は11.025KHz、8ms毎のフレーム周期で分析
- ◎特徴パラメータはメルケプストラム係数
- ◎110音節の連続出力分布型5状態HMM
- ◎男性36名の文音声による不特定話者のモデル
- ◎認識方法はオートマトン制御One Pass DP法

今回の実験のために、オートマトン付きOne Pass DP認識プログラムを作成した。このプログラムは、オートマトンに基づいて、タスクの辞書およびオートマトン情報ファイルを作成すれば、どんな制御でも動くようになっている。不特定話者の音声認識結果としては、上位二つの候補を出力するようになっている。

2.2 ファイルの記述

辞書ファイルは、単語シンボル、単語名、音節辞書の三つの部分から構成されている。単語シンボルは自由に決められる。

オートマトン情報ファイルは、オートマトン制御図に基づいて、最終状態の集合を1行目に記入する。2行目以後は各状態の番号と行き先の状態番号及びその状態遷移に伴う全ての単語シンボルのリストを記述する。行き先状態番号に遷移してくる状態がない場合は、状態番号の所に-1を記入する。

A Robot Simulation System
with Speech Input and Touch Input
JianXin Zhang, Seiichi Nakagawa
Toyohashi University of Technology
1-1, Tenpaku, Toyohashi, Aichi 441, Japan

2.3 例

図1のような簡単なロボット制御用オートマトンに対して、辞書ファイルとオートマトン情報ファイルをそれぞれリスト1とリスト2に示した。

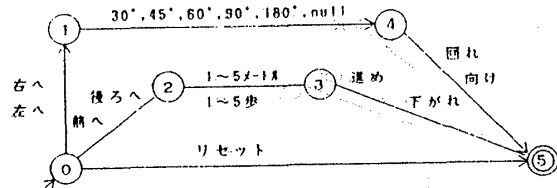


図1. 簡単なオートマトン制御例

リスト1:

12	←	単語数	
maehe		前へ	ma e e
usirohe		後ろへ	u si ro e
migihe		右へ	m i gi e
hidarihe		左へ	hi da ri e
1m		1メートル	i ti me e to ru
2m		2メートル	ni me e to ru
susume		進め	s u su me
sagare		下がれ	sa ga re
maware		回れ	ma wa re
muke		向け	m u ke
reset		リセット	ri se to
null		NULL	

↑ ↑ ↑
単語シンボル 結果表示用発音辞書

リスト2:

5	←	最終状態の集合
-1		0
0		1 migihe hidarihe
0		2 maehe usirohe
0		5 reset
1		4 30° 45° null
2		3 1m 2m
3		5 susume sagare
4		5 maware muke

↑ ↑ ↑ ↑
| 行き先状態番号 |
状態番号 単語シンボル

3・シミュレーションシステム

今回は、音声入力のみ、タッチ入力のみ、両方の組合せによる簡単なロボット制御システムを作成した。状況設定としては、ロボットの頭にカメラを付けて、周囲の状況を画面に表示し、人間が画面を見ながら、次の三つのシミュレーションシステムでロボットをコントロールし、未知の地域の探検や標本収集、機械故障修理などを行なうことを想定している。

ロボット動作の設定は、①ロボットの視野範囲は、アームを中心にして、±100度の範囲とする。②ロボット自身は回転、移動ができるとする。③ロボットは、物体を収集することや故障機械のネジを外すこと、修理用機械のボタンを押すことなどもできるとする。

以上の設定に基づいて、次の三つのシステムを作成した。

3. 1・音声とタッチ入力システム

オートマTON制御図(図2)によって、約千種類のコマンドが音声とタッチ入力(静電式、読み取り誤差:8.8mm×6.3mm)で生成できる。このシステムは、音声とタッチが同時に入力でき、しかも相互補助もできる。例えば、音声認識の結果によって、タッチ入力を修正する例としては、音声認識の結果が「この三角を取れ」となった場合、タッチ入力の座標を強制的に一番近い三角の座標に修正する。また、タッチ入力によって、音声認識結果を修正する例としては、音声認識の結果が「右へ30度回れ」の場合は、タッチ入力のない時は、そのまま動作するが、タッチ入力があった時に、その入力が認識結果の回転方向と矛盾すれば、認識結果を修正する。

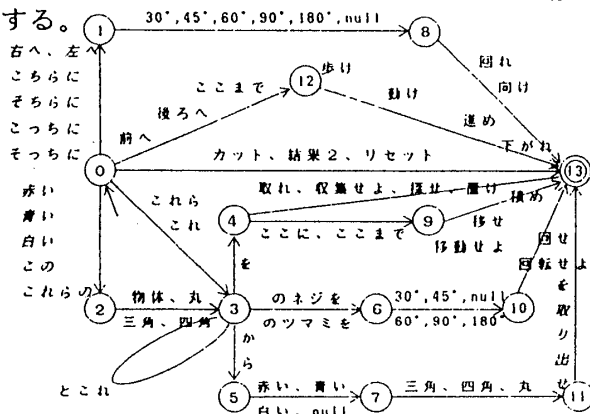


図2・音声認識とタッチパネルによるロボット制御のオートマTON図(語彙:51)

3. 2・音声入力システム

図2中のタッチ入力に関する単語(この、ここに、など)を外して、音声入力だけで約360種類のコマンドを生成することができる。

3. 3・タッチ入力システム

ロボットの表示部分とメニュー(32個)の箱にタッチすることによって、ロボットの動作を制御する。

4・実験

最初の実験は、決められたコマンドを被験者に操作させて、ロボットの動作を確認する。次の実験は、ある目的を被験者に与えて、被験者が考えたコマンドを使ってその目的を達成する。最後に三つのシステムについてのアンケートに答えてもらって実験を終了する。

5・評価

表1と表2に、それぞれ各システムの特徴と被験者のアンケートによる評価の結果を示している。その他に、「もし音声認識率が100%で、しかもリアルタイムでできたら、どのシステムが一番使いやすいか?」という質問に対して、殆どの被験者は、「組合せ入力」と解答した。

表1・各システムの特徴

項目	入力方法	音声入力	タッチ入力	組合せ入力
システム語彙数		多くできる	多くできない	多くできる
見えない物体指定		できる	できない	できる
重なる物体の指定		できる	できない	できる
特徴のない物体の指定		できない	できる	できる

表2・実験による評価結果

項目	入力方法	音声入力	タッチ入力	組合せ入力
速度		遅い	早い	やや早い
認識精度		わるい	ややよい	ややわるい
習得難易度		やや難しい	易しい	難しい
離人差、雑音の影響		あり	なし	あり
入力の自由度*		多い	少ない	多い

*何をやるか、どうやるか、どのボタンをどのように押すかなど

6・まとめ

音声とタッチ入力を組み合わせることによって、タッチ入力の欠点をほぼなくすることができた。しかし、音声とタッチ入力のタイミングや、使い方が複雑になるなどの問題点を生じた。これから、より高速かつ正確・柔軟なマルチモーダルインターフェースを目指して研究を進める予定である。

参考文献:

- [1]安藤他; 音声とポインティングジェスチャを入力手段とする「インテリアデザイン支援システム」の試作と操作性の評価, 情報処理学会, 音声言語研究グループ資料, 93-SPL-2-2
- [2]谷正之; 産業システムにおけるヒューマンインターフェース技術の一動向, 電学誌, 113巻, 5月号, 1993
- [3]新田他; 自由発話音声入力と直指を利用したマルチモーダル対話システムの検討, 信学技報, SP92-120(1993-01)