

スマートスピーカーを利用したロボットコントローラーの設計

S16022 磯井 大輝

1. はじめに

本研究は、スマートスピーカーを利用して、ロボットとの対話を通じて、ロボットを操作するシステムの構築が目的である。これまでの研究では、二足歩行ロボットに対しての会話の作成を行なった。

本研究では、二足歩行ロボットに対して、サッカーを実行させることを目的として、歩行やキックなどの 11 種類の音声コマンドの登録を行ない、実際のロボットの制御実験を実施した。

2. スマートスピーカーとロボット



図 1 Google home mini と i-SOBOT

図 1 は、本研究で利用するハードウェアを示した。スマートスピーカーには Google 社製の Google Home mini を利用し、二足歩行ロボットにはタカラトミー社製の i-SOBOT を利用した。また、ロボットを制御するパソコンから i-SOBOT のインターフェースとして、BUFFALO 社製の Remote Station を使用した。

3. 音声コマンド

音声コマンドでは表 1 のように「前進」から「終了」までの 11 種類のコマンドを作成し、登録を行なった。また、「前進」から「右旋回」までは、歩

数を表す変数を設定しており、例えば「前進 4 歩」や「右に 2 歩」など、歩数を指定することが可能である。

表 1 音声コマンド

	音声コマンド	変数
1	前進	〇〇歩
2	後退	〇〇歩
3	左	〇〇歩
4	右	〇〇歩
5	左旋回	〇〇歩
6	右旋回	〇〇歩
7	左キック	
8	右キック	
9	立つ	
10	逆立ち	
11	終了	

4. 実験

本研究では、音声コマンドを発言して Google Home が音声認識を行いロボットが動作を開始するまでの時間と、ロボットの動作が完了し Google Home が返答するまでの時間を計測した。

3 人の被験者に対して 10 回ずつコマンドを実行してもらった。それぞれの被験者は、Google Home に対して「前進〇〇歩」という音声コマンドを実行した。このコマンドでは Google Home は「前進」というコマンドの種類と「〇〇歩」といった歩数に関する変数の両方を認識しなければならない。

5. 実験結果

表 2 から表 4 にそれぞれの被験者の結果をまとめた。R はロボットが動作を開始するまでの時間であり、G は Google Home が返答するまでの時間を示している。認識できなかった項目は×を記

している。また、歩数が認識できなかった場合は、プログラム上 2 歩進むという仕様としている。また、「前進 4 歩」のように歩を言う場合と、「前進 4」のように歩を言わない場合の結果を示した。

表 2 被験者 A

被験者A	歩あり		歩数	歩なし		歩数
	R	G		R	G	
1	3.65	5.36	5	6.08	11.31	5
2	3.56	6.61	×	×	×	×
3	×	×	×	×	×	×
4	×	×	×	3.83	9.38	7
5	5.7	7.93	1	3.53	5.55	1
6	3.66	8.68	6	3.51	10.16	6
7	3.16	5.96	×	3.83	7.41	4
8	3.7	6.58	×	3.14	12.04	9
9	3.26	11.6	8	3.55	11.97	8
10	3.2	14.41	10	5.9	8.98	×
平均	3.736	8.391		4.171	9.60	

表 3 被験者 B

被験者B	歩あり		歩数	歩なし		歩数
	R	G		R	G	
1	×	×	×	4.65	9.61	5
2	3.1	5.73	×	3.88	7.5	3
3	4.05	6.56	×	3.5	5.71	1
4	3.53	6	×	3.58	8.27	4
5	2.98	10.43	7	3.31	10.75	7
6	3.93	12.03	8	2.84	11.33	8
7	×	×	×	×	×	×
8	3.36	6.47	×	3.24	13.03	9
9	×	×	×	×	×	×
10	3.56	6.17	2	5.51	8.41	×
平均	3.501	7.627		3.814	9.326	

表 4 被験者 C

被験者C	歩あり		歩数	歩なし		歩数
	R	G		R	G	
1	3.55	7.61	5	6.52	9.62	5
2	×	×	×	6.7	10.08	5
3	4.11	6.74	×	5.1	11.17	7
4	3.6	9.87	7	4.11	9.95	7
5	3.21	6.12	×	3.43	6.03	×
6	3.95	10.38	6	3.98	10.55	6
7	3.7	6.81	×	3.96	6.77	×
8	3.8	6.53	3	3.86	6.55	3
9	3.55	6.43	×	3.5	12.01	9
10	3.13	6.05	×	4.16	13.27	9
平均	3.622	7.393		4.532	9.60	

まず、それぞれの被験者のコマンドそのものの

認識率は、被験者 A で 80%(歩ありで 10 回中 8 回, 歩なし 10 回中 8 回), 被験者 B で 75%(歩ありで 10 回中 7 回, 歩なし 10 回中 8 回), 被験者 C で 95%(歩ありで 10 回中 9 回, 歩なし 10 回中 10 回)の認識率となった。被験者 A は声がやや小さく, Google Home から顔がやや遠くの位置にあったためコマンドの認識が 80%となったと考えられる。被験者 C では声が大きく Google Home との距離が近くであったためコマンドの認識が 90%となった。

また、歩数の認識において、「歩あり」の場合には、被験者 A で 50%, 被験者 B で 30%, 被験者 C で 40%であった。それに対して、「歩なし」の場合には、被験者 A で 70%, 被験者 B で 70%, 被験者 C で 80%であり、「歩あり」の場合に比べて、高い認識率であった。被験者の数字の言い方にも違いがあり、「前進 6 歩」に対して「ろくほ」や「ろっぽ」といった違いで Google Home の認識が変わった。Google Home に数字を認識させるだけなら「歩」を言わない方が認識は高くなることがわかった。

R の全体の平均は 3.896 秒であり、音声コマンドを実行して、ロボットが動作を開始するまでに、およそ 4 秒程度必要であることがわかった。

6. おわりに

本研究では、二足歩行ロボットに対して、サッカーを実行させることを目的として、歩行やキックなどの音声コマンドの登録を行ない、実際のロボットの制御実験を実施した。

その結果、スマートスピーカーに近い位置でより大きな声で発言した方が認識率は高くなり、歩数などの数を認識させる場合には、単に数値のみを発言した方がいいことがわかった。また、ロボットが動作するまでには、およそ 4 秒程度の時間が必要であることがわかった。

参考文献

1) Action on Google

<https://developers.google.com/actions/>