

スマートスピーカーを利用した口 ロボットコントローラーの設計

平石研究室
S16022 磯井大輝

はじめに

- スマートスピーカーを利用して, ロボットの操作をするシステムの構築
⇒Google Homeを対象
- 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
⇒11種類のコマンド登録とロボットの制御実験を実施

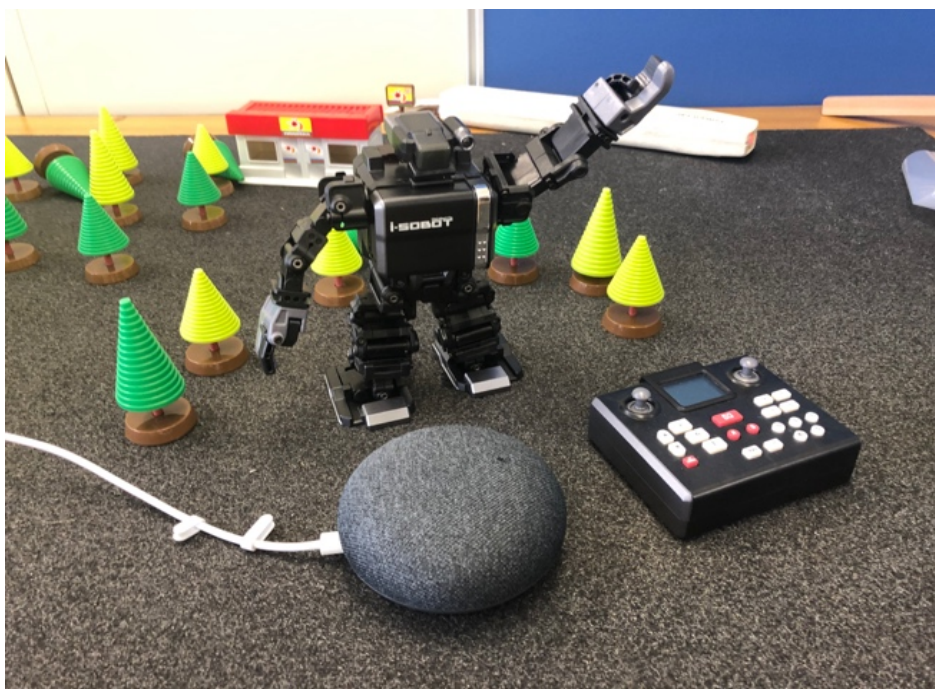
前期のまとめ

- アプリケーションの使い方
- 会話の作成
- ロボットサーバーの作成

後期の課題

- 会話の種類登録
- ロボットコマンドの送信部分の作成
- 音声認識の実験

スマートスピーカーとロボット



- スマートスピーカー

Google社製のGoogle Home mini

- 二足歩行ロボット

タカラトミー社製のi-SOBOT

- Remote Station

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

会話の種類

- 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た.
- 「前進」から「右旋回」まで変数を設定している.

例

「前進4歩」や「右旋回5歩」など

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	

実験

- 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(R)と, ロボットの動作が完了しGoogle Homeが返答するまでの時間(G)を計測.
- 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.
- 3人の被験者に対して10回ずつコマンドを実行. 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.
- 「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.

実験結果

表2 被験者A

被験者A

- コマンド認識率 80%

⇒歩ありで10回中8回, 歩なしで10回中8回

- 歩数の認識率

⇒歩あり 50%, 歩なし 70%

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	

実験結果

表3 被験者B

被験者B

- コマンド認識率 75%

⇒歩ありで10回中7回, 歩なしで10回中8回

- 歩数の認識率

⇒歩あり 30%, 歩なし 70%

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	

実験結果

被験者C

- コマンド認識率 95%

⇒歩ありで10回中9回, 歩なしで10回中10回

- 歩数の認識率

⇒歩あり 40%, 歩なし 80%

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	

実験結果

- 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
- 声の大きさ，顔とGoogle Homeの位置
⇒被験者ごとに違いがあった
- 歩ありの数字の言い方
⇒「6歩」に対して「ろくほ」や「ろっぽ」の違い

動画



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。