

Raspberry Pi と Arduino を用いたソーシャルロボットに関する研究

A Study on social robot using Raspberry Pi and Arduino

小比田 尚紀¹⁾, 加登 大地¹⁾

指導教員 林 誠治¹⁾

1) 拓殖大学工学部 電子システム工学科 林研究室

キーワード：ソーシャルロボット、遠隔操作、音声合成、音声認識、雑談会話

1. はじめに

これからの世の中では身の回りにロボットが存在するのが当たり前になり、ロボットと人がコミュニケーションをとる機会が増えると考えられる。本研究ではマグボット[1]をベースに音声認識機能を加え、人とのコミュニケーションをより自然に行うことを目的とする。

2. 音声認識を搭載したマグボット

2.1. マグボットについて

マグボット(図 1)はオープンソースのロボットで Arduino と Raspberry Pi(図 2) を組み合わせて作られている。Raspberry Pi の中には Web サーバーを設置しており、PC・スマートフォン・タブレットから Web ブラウザ経由でアクセスし遠隔操作が可能である(図 3)。



図 1: 本研究中で開発中のマグボット

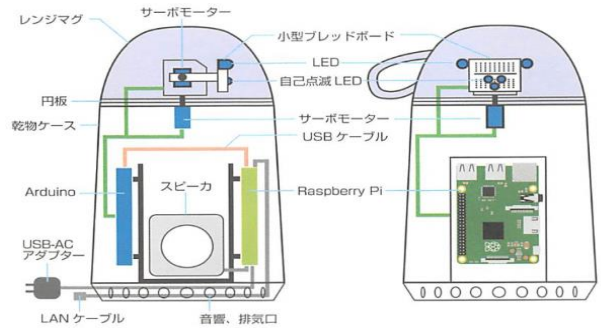


図 4: マグボットのハードウェア[1]

表1 スピーカの仕様

部品	型番
アンプ	PAM8012
スピーカ	FA200C
電源	3V

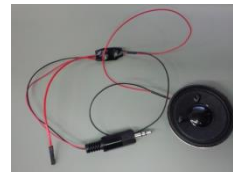


図 5: アンプ付スピーカ

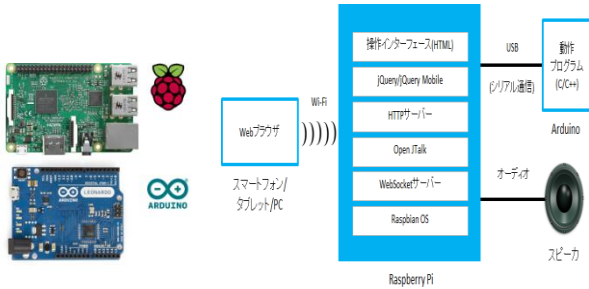


図 2: Raspberry Pi と Arduino

図 3: 通信システム

2.2. ハードウェア

2.2.1. ボット本体

Arduino はハード制御を、Raspberry Pi はネットコミュニケーションと合成音声・音声認識・雑談会話を担当する。また電力は Raspberry Pi から USB ケーブルで Arduino に供給し、同時にシリアル通信を行う(図 4)。

2.2.2. アンプ付スピーカ

小型のアンプ付スピーカを製作し、これを使用することで内部の省スペース化を可能にした。スピーカの仕様を表 1 に、スピーカを図 5 にそれぞれ示す。

2.3. ソフトウェア

2.3.1. HTTP サーバー

HTTP サーバーは Nginx(エンジンエックス)を使用する。Nginx は少ないメモリで高速に動く軽量サーバーである。図 6 に示すように PC、スマートフォン、タブレット等の Web ブラウザから有線/無線 LAN 経由で Raspberry Pi 内の HTTP サーバー にアクセスし、マグボットの操作インタフェースである HTML ファイルを開くことができる。

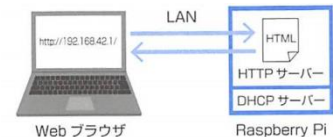


図 6: HTTP サーバーの使用用途[1]

2.3.2. WebSocket サーバー

WebSocket サーバーはブラウザとサーバーの間で双方向通信、非同期通信を実行するためのプロトコルを使用したサーバーである。マグボットでは図 7 のように操作インタフェースから WebSocket サーバーを経由して Arduino に接続した LED や、サーボを制御するために用いる。

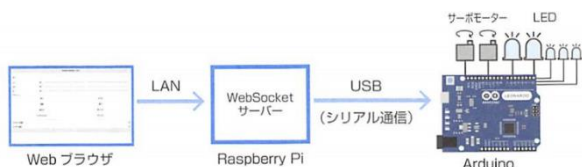


図 7: WebSocket サーバーの使用用途[1]

2.3.3. 日本語音声合成ソフトウェア Open JTalk

Open JTalk とは名古屋工業大学の徳田・李・南角研究室が開発している文字列を音声(Wav など)に変換するオープンソースの日本語テキスト音声合成システムである。マグボットでは図 8 のように操作インタフェースから文字列を入力して Open JTalk によりその Wav データを取得し、`aplay` によりスピーカから出力する。

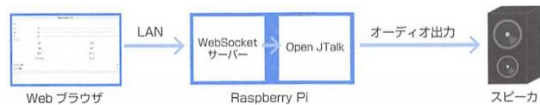


図 8: Open JTalk の使用用途[1]

2.3.4. JavaScript ライブラリ jQuery

jQuery は Web ブラウザ用の JavaScript コードをより容易に記述できるようにするために設計された軽量な JavaScript ライブラリである。そのため少ない記述でコード書くことができる。

2.3.5. Julius 音声認識

Julius は音声認識システムの開発・研究のためのオープンソース大語彙連続音声認識エンジンである。単語辞書のモジュールを組み替えることで、ソーシャルロボットに特化した音声認識語彙システムを構築できる。本稿では、音声の収録に USB マイクを使用し、Raspberry Pi 上に Julius 音声認識ソフトウェア(Julius 本体、ディクテーションキット、グラマーキット)を実装した[2]。

また、認識したい言葉として「笑って・悲しい・びっくり・ウィンク・ハイテンション・どこから来たの・終わり」を単語辞書に登録し、ディクテーションファイル(口述筆記)を生成した後、Julius ディクテーション実行キットを使用して Julius 音声認識を実行した。Python スクリプト上で、モジュールモードの Julius サーバーをバックグラウンドで動作させ、`socket` モジュールを用いて Julius からの認識結果を取得できるようにした。Julius サーバーの受信データから RECOGOUT キーを XML として随時パースし、単語辞書中の言葉を判別して、各々の言葉に対する動作(サーボ動作、LED 制御、音声合成)を行っている。

2.3.6. Google Speech API

前項で述べた Julius はオフラインでの動作であるが、インターネットクラウドサービスを利用するオンライン音声認識では一般に認識精度の向上が見込める。ここでは Google Speech API のオンライン音声認識を使用する[3]。API を利用するには Web 上から Google API キーを取得し、`requests` ライブラリの `post()` メソッドを使

用して、認識用 URL と Wav 形式の音声データ、サンプリングレートなどを指定する。認識結果は JSON(JavaScript Object Notation)形式で返ってくるので、音声認識の結果候補の中から最初の `transcript` の要素(信頼性が最も高い結果)を本研究での認識結果として採用した。

2.3.7. docomo 雑談会話 API

ロボットと人間との自然な会話を実現する方法として、ここではオンラインの docomo 雑談会話 API を利用する[4]。API を利用するには Web 上から docomo API キーを取得し、`requests` ライブラリの `post()` メソッドを使用して、雑談会話用 URL とこちらから話しかける言葉などを JSON 形式のペイロードとして指定する。ペイロードにはニックネーム、性別、年齢、出身地など話者の情報を指定できる。会話が成功すると JSON 形式のオブジェクトが返ってくるので、その `utt` の要素を取り出すことで docomo からの雑談会話の返事を受け取ることができる。

3. システムの動作例

システムの動作例として、マグボットの「笑う」の動作について説明する。まず、Raspberry Pi 内の HTTP サーバーにアクセスし、HTML ファイルを Web ブラウザに表示する。次に操作インタフェースの「笑う」ボタンをクリックすることにより、WebSocket 通信を利用して Raspberry Pi に“@L”の文字列を送る。“@”から始まる文字列は Arduino の制御(動作させるコマンド)として認識させているので、そのままシリアル通信を利用して Arduino に送り、対応するボットの顔に相当する LED と上下のサーボモーターを動作させることで「笑う」の動きを実現する(図 9)。

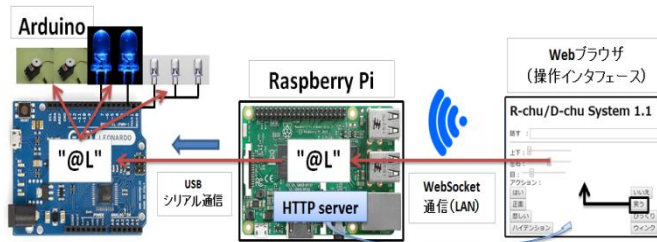


図 9: システムの動作例

4. 今後の予定

今後の予定としては、人感センサー、カメラモジュール、自走機構などの新たな機能を組み込み、引き続き改良していきたいと考えている。

参考文献

- [1] 小池星多, “おしゃべりロボット「マグボット」ラズパイと Arduino で電子工作”, リックテレコム, 2016 年
- [2] “大語彙連続音声認識エンジン Julius”, <http://julius.osdn.jp/>
- [3] “Speech API-音声認識”, <https://cloud.google.com/speech/?hl=ja>
- [4] “docomo Developer support”, <https://dev.smt.docomo.ne.jp/?p=index>