

論文

小型漫才ロボットの制御システムの開発

孝橋 一希^a, 真下 遼^b, 久井 晶太^a, 梅谷 智弘^a, 北村 達也^a, 灘本 明代^a

^a 甲南大学 知能情報学部 知能情報学科

神戸市東灘区岡本 8-9-1, 658-8501

^b 甲南大学大学院 自然科学研究科 知能情報学専攻

神戸市東灘区岡本 8-9-1, 658-8501

(受理日 2015 年 5 月 14 日)

概要

2 台のロボットによる漫才システムを小型ロボットで実装するために, 新たな制御システムを RT ミドルウェア上で使用可能なコンポーネントとして開発した. この制御システムは, 漫才台本自動生成システムで生成される XML ファイル形式の漫才台本をそのまま使用でき, 音声合成ソフト AITalk を用いて生成された関西方言の合成音声による漫才ができる.

キーワード: 漫才ロボット, RT ミドルウェア, 関西方言音声合成

1 はじめに

甲南大学知能情報学部で開発されたコミュニケーションロボットは, インターネット上のウェブニュース, ビッグデータを用いて, 与えたお題に応じた漫才の台本を自動で生成することができ, 2 台のロボットを用いた掛け合い漫才をすることができる [1]-[4]. これらのロボットは, 全高 1 m のあいちゃん, 0.5 m のゴン太から成り, 有機 EL ディスプレイの眼によって表情を表示しながら合成音声による漫才の対話を実行する.

この 2 台のロボットによる掛け合い漫才は新聞, テレビなどのメディアで紹介され, オープンキャンパスや外部のイベントなどで展示が行われた. しかし, ロボットが大型のため運搬に費用がかかることや, プログラムが複雑な構成で改良が困難であることが問題になった.

そこで, 自動生成された漫才台本を用いて大型の漫才ロボットと同様の漫才対話を実現する制御システムおよび, 小型で運用しやすい 2 台の漫才ロボットを開発した. 本稿では, この小型漫才ロボットの制御システムについて述べる.



図 1: 小型漫才ロボット (左: あいちゃん, 右: ゴン太)

2 小型漫才ロボット

小型漫才ロボットは高さ 25 cm のあいちゃん (図 1 左), 高さ 15 cm のゴン太 (図 1 右) の 2 台から成り, それぞれに PC を 1 台ずつ用いて制御を行う. PC の OS は Windows 7 である. 1 台目の PC は次の役割を担う (図 2).

1. 漫才台本の生成
2. 表情データの表示用 WWW サーバ (Apache)
3. 漫才セリフ音声の合成
4. 漫才台本の進行制御
5. 1 台目のロボット (あいちゃん) の動作制御と音声再生
6. 2 台のロボットの表情データ書き換え処理

もう 1 台の PC では, 2 台目のロボット (ゴン太) の動作制御と音声再生を行う. ロボット台車の制御, 音声再生はそれぞれの PC に分担させることによって, 音声の発生源の変更やハードウェア管理を容易にしている [5].

ロボットの台車部分にはヴイストン社の Beauto Rover RTC-BT, 顔部分に Apple 社の iPod touch をそれぞれ 1 台ずつ用いている. 顔表情は iPod touch 付属のブラウザ Safari で WWW サーバ上の Web ページを参照することによって提示する. 顔表情は PNG 形式のファイルで iPod Touch の画面サイズに合わせて作成した (図 3). 表情の更新は WWW サーバで再生するファイルを差し替え, 表示する Web ページを毎秒再読み込みすることによって実現した. 音声ファイルは, エーアイ社の AITalk 関西弁話者 SDK を用いて生成しており, Windows のファイル共有によりそれぞれの PC からアクセスする.

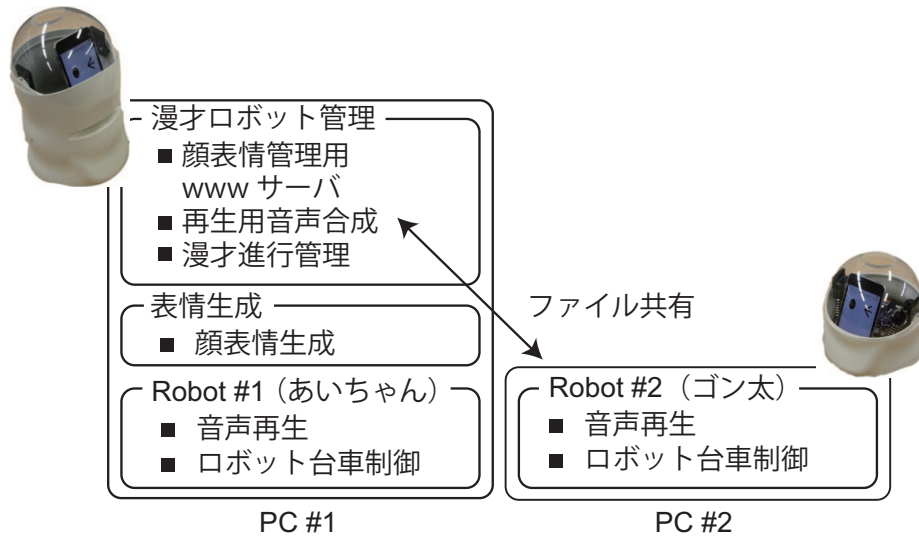


図 2: 小型漫才ロボットのシステム構成

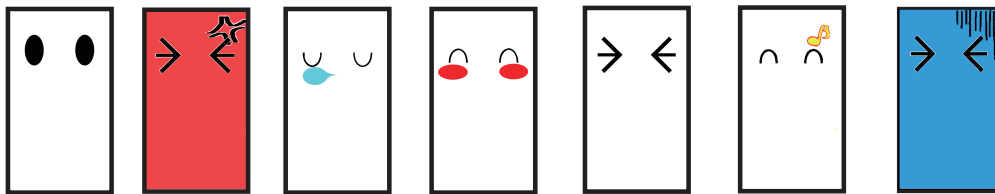


図 3: 小型漫才ロボットの顔表情例

このシステム構成をもとに、それぞれの機能を RT ミドルウェア上で利用する RT コンポーネントとして実装した。RT ミドルウェアについては後述する。2 台の PC の OS は Windows 7, 開発環境として Open RTM-aist 1.1.0 の C++ 版, 同 Python 版と Visual C++ を使用した。作成した RT コンポーネントは、以下の 4 種類である。

1. 漫才ロボットシステム制御
2. 漫才演出時の顔表情生成
3. ロボット台車の制御
4. 音声再生

このうち、2, 3, 4 はそれぞれのロボットに対してコンポーネントを割り当てて制御する。

```
## AITalk2SDK config ##  
AUTH: C:%AI%AITalkIII%license%aitalk_win.lic  
DBPATH: C:%AI%AITalkIII%dic%voice  
DBRATE: 22050  
RCPATH: C:%AI%AITalkIII%dic%lang%kansai  
VOICE: miyabi_22  
WAVERATE: 22050  
MULAW: 0
```

図 4: config_mary.ini

```
## AITalk2SDK TTS parameter ##  
PITCH: 1.15  
RANGE: 1.25  
SPEED: 0.9  
VOLUME: 1.00  
  
# BREAK in a sentence.  
PAUSE1 : 150  
PAUSE2: 250  
PAUSE3: 400
```

図 5: tts_mary.ini

3 音声合成について

本システムではエーアイ社の AITalk 関西弁話者 SDK を用いて音声合成を行った。この音声合成 SDK には、女性話者“みやび”，男性話者“やまと”の 2 人の話者が提供されている。本システムでは前者をあいちゃん，後者をゴン太に割り当てた。

音声合成には、発話内容が記述されたテキストファイル，話者や合成音声の抑揚，速さなどのパラメータが記述された設定ファイル 2 つが必要である。例としてあいちゃん用の 2 つの設定ファイルを示す。config_mary.ini (図 4) は各種ファイルのパス，音声データの形式，話者について設定し，tts_mary.ini (図 5) は合成音声の基本周波数，その変動幅，話速，音量，ポーズ長について設定する。設定ファイルの各パラメータは，複数のパラメータセットで合成した音声を聴き比べ，各ロボットのキャラクタに適合していると感じられた値を設定した。

漫才に用いる音声を生成するために，SDK 付属のサンプルプログラムを改良し，漫才中のセリフが記述されたテキストファイルと各ロボットごとに異なる設定ファイルのパスを引数として与えて WAV ファイルを生成できるようにした。このプログラムを後述するプログラムから呼び出し実行す

ることによって音声ファイルを生成する。

なお、漫才実行中に一つ一つのセリフを合成すると、計算に時間を要し、スムーズな対話が困難となるため、音声は漫才実行前にまとめて合成する。

4 漫才台本の自動生成について

本章では、漫才台本の自動生成および生成される台本の形式について概説する。

4.1 漫才台本自動生成システム

この漫才台本生成システムは、人名や一般名詞などのお題を与えると、インターネット上からお題に関連するニュース記事を検索し、その記事を元につかみ、本ネタ、オチから成る漫才台本を生成する。本ネタの中では読みの似た言葉の言い間違いや、言葉の印象語でボケに同調してから訂正するノリツッコミ、認知度が近い単語と間違えさせる対立ボケなどのネタを生成し、オチではお題に応じたなぞかけを生成する。このシステムによって生成された漫才台本はXMLファイル形式で出力される。

4.2 生成された台本の形式

XMLファイル形式で出力される漫才台本は、各行がひとつの命令文となっている。命令文はロボットの動作、表現、台詞の3種類に分けられる。以下にその3種類の例を記す。台本中のMary, Bobはそれぞれあいちゃん、ゴン太に対応している。

- `<look name="Mary", what="audience"/>`
漫才中にロボットがどこを向くかについての情報で、Mary, Bobの“どちら”が、“どこ”が向くかを指定している。この例では、“Mary”が“audience”（観客側、正面）を向くように指定している。
- `<PEmo name="Bob">PE11/>`
漫才中のロボットの表情についての情報で、Mary, Bobの“どちら”が、“どの表情”をするかを指定している。“PE〇〇”の番号で表情の指定をして、番号に対応した顔表情を表示する。“〇〇”に含まれる番号は00から49までの50種類の2桁の数字である。
- `<cast name="Mary"> どもー、あいちゃんです </cast>`
漫才のセリフの情報で、Mary, Bobの“どちら”が話者であるかと、セリフの内容を指定している。

5 小型漫才ロボットのプログラム

本章では、小型漫才ロボット用に開発したプログラムについて説明する。プログラムはRT ミドルウェア上で使用するRT コンポーネントとRT コンポーネントから関数として呼び出し使用されるプログラムに大別される。前者はPython およびC++で、後者はPython で開発した。

RT ミドルウェアとは分散型ロボット用ミドルウェアの一つで、ロボット用ソフトウェアをモジュール化したRT コンポーネント群をポートと呼ばれるインターフェースを介して通信を行うことによってロボットシステムを構築するための通信規格である [6]。これを利用して開発したことによって、ソフトウェアの再利用性を高めるとともに、短期間での開発を実現した。また、RT ミドルウェア上で開発したため、他のハードウェアへの移植性も向上した。

5.1 小型漫才ロボットの Python プログラム

RT コンポーネントから呼び出されるプログラムを以下に示す。

- `mkmanzai_AI.py`

漫才台本ファイルから漫才実行に必要なファイル一式を生成するプログラムである。XML 形式の漫才台本ファイルのパスを引数として実行する。

このプログラムでは、まず XML ファイルを一行ずつ読み込み、行中に“cast”を含む場合、話者名、セリフ部分を抜き出し、セリフをテキストファイルとして書き出す。

次に、上記のテキストファイルと音声合成の設定ファイル（話者名によって指定される）からセリフの音声を合成する。これによって生成された音声ファイルは、話者と台本中でのセリフの順番によって、“(話者名)(番号).wav”と命名され、音声ファイル用フォルダに保存する。同時に、セリフのテキストを音声ファイルに対応した“(話者名)(番号).txt”というファイル名で漫才のセリフ用フォルダに保存する。

最後に、上記の XML ファイル中の“cast”を含む行を音声ファイルのパスに書き換えたものをテキストファイルとして保存する。このファイルは漫才の実行時に読み込み使用する。

- `searchfiles.py`

フォルダと文字列を引数として実行すると、指定したフォルダ中から指定した文字列をファイル名に含むファイルを検索して一覧表示するプログラムである。

本システムでは指定したフォルダ中にある漫才台本ファイルを番号付きの一覧として表示し、実行する漫才台本ファイルの選択を番号の入力によって補助する役割を担う。

- `changeEyes_mary.py`, `changeEyes_bob.py`

PE と二桁の数字から成る“PE01”などの顔表情番号を引数として、顔表情画像フォルダ内の表示用顔画像のファイルを番号に対応した顔表情画像に変更するプログラムである。

プログラム中に顔表情番号とそれに対応する顔表情のパスを記述しており、それぞれのロボットに対応させ、ロボットごとに異なる表情を表示することができる。

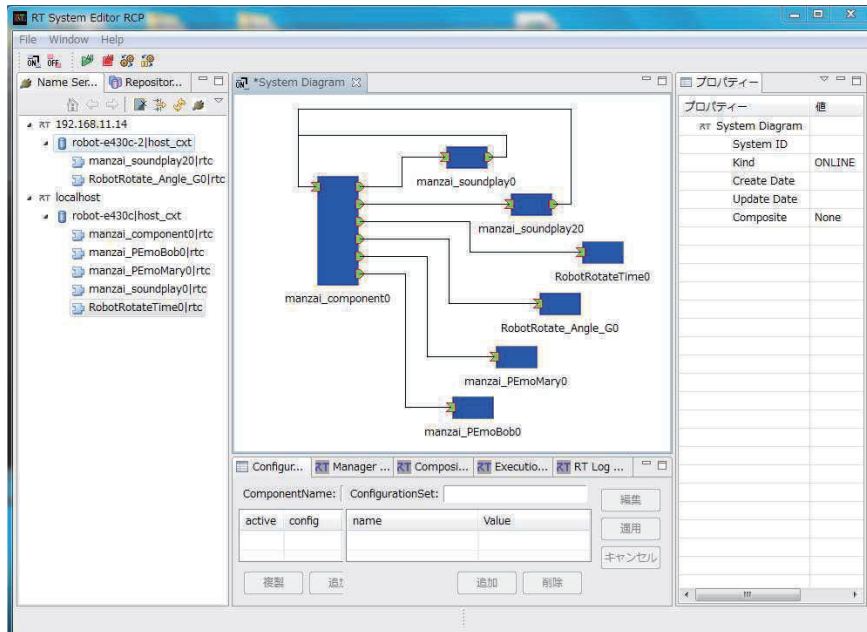


図 6: RTSystemEditor 上でのコンポーネント接続図

5.2 RT コンポーネントの構成

RTSystemEditor 上で起動する RT コンポーネントは以下のもので構成される。図 6 に各コンポーネントを接続した図を示す。

- manzai_component

2 台のロボットに台本の流れに沿った動作をさせるための制御を行うコンポーネントである。コマンドラインに表示される漫才台本一覧の番号などを入力することによって、漫才台本から音声ファイルの生成、漫才の実行開始ができる。このコンポーネントは、

1. 再生する音声ファイルのパス
2. 1つ前に再生した音声ファイルのパス
3. 現在ロボットが向いている方向
4. 表示する顔表情の番号

の 4 つのデータを入出力用のポートによって他のコンポーネントへと受け渡す。

これらのコンポーネントは、入出力用のポートとして入力には音声ファイルの再生状況を取得するためのポートを 1 つ持ち、出力には再生する音声ファイルのパスを送るためのポート、台車の回転量を送るポート、顔表情の番号を送るポートをそれぞれ 2 つずつ持つ。

このコンポーネントを起動すると、まずセリフの音声合成するかを Y/N で確認する。音声合成を行わない場合 (N) は、最後に実行した漫才台本を実行する。音声合成を行う場合 (Y) は、漫才台本の番号付きリストが表示され、使用する台本を番号で指定すると音声合成を開始する。

この処理の終了後、Enter キーを押すと漫才を実行できる。このときに秒単位で数値を入力すると、漫才の開始を指定した時間だけ遅らせることができる。

漫才を開始すると、XML ファイルの書き換えたテキストファイルを一行ずつ読み込んで、その内容に応じた処理を実行する。

1. 読み込んだ行に音声ファイルの拡張子 (.wav) が含まれる場合、音声再生コンポーネントに音声ファイルのパスを引き渡す。その後、音声再生コンポーネントの出力信号を監視し、それが音声再生終了の信号に変わると、次の処理に移る。
2. 読み込んだ行に“look”が含まれる場合、動作制御コンポーネントにロボットの回転量を引き渡す。
3. 読み込んだ行に“PEmo”が含まれる場合、顔表情制御コンポーネントに顔表情の番号を引き渡す。

漫才が終了するとコンポーネント起動時の処理に戻る。

- manzai_soundplay, manzai_soundplay2

音声を再生するコンポーネントであり、それぞれのロボットに1つずつ用意して使用する。入出力用ポートとして、入力には音声ファイルのパスを受け取るポートを1つ持ち、出力用には音声の再生状況を出力するポートを1つ持つ。

manzai_component から音声ファイルのパスの読み込みを繰り返し、直前に再生した音声ファイルのパスと異なる場合に音声を再生する。音声の再生が終了すると、manzai_component に音声の再生終了の信号を出力する。

- RobotRotateTime, RobotRotate_Angle_G

ロボットの回転動作を制御するコンポーネントであり、それぞれのロボットに1つずつ用意して使用する。入力用ポートとして台車の回転量を受け取るポートを1つ持つ。

manzai_component から出力された回転量を読み込み、ロボットの台車に回転動作を行わせる。

- manzai_PEmoMary, manzai_PEmoBob

顔表情を変更するコンポーネントであり、それぞれのロボットに1つずつ用意して使用する。入力用ポートとして、顔表情番号を受け取るポートを1つ持つ。

manzai_component から出力された 顔表情番号を読み込み、それを changeEyes_mary または changeEyes_bob (5.1 節参照) に引数として与えることによって顔表情を変更する。

6 漫才の実行方法について

本章では、前章で説明したプログラムを使用して漫才を実行する方法を説明する。ここで、漫才台本自動生成システムにより生成された漫才台本は、manzai_component が参照するフォルダに予め移動させておいたものとする。

6.1 通常の実行方法

まず、あいちゃん側の PC で RTSystemEditor を起動する。また、顔表情を表示させるために WWW サーバを立ち上げる。

次に、各 PC でネームサーバ、RT コンポーネントを起動し、RTSystemEditor 上で接続する。このとき、ファイアウォールが有効であると通信ができないため一時的に無効にする。

最後に、接続したコンポーネントをすべて Activated にする。そして、manzai_component のウィンドウに表示されるメッセージに従い入力を行うことによって音声ファイルの生成や漫才の実行ができる。音声ファイルを新たに生成しない場合は最後に実行した漫才を実行する。

6.2 保存されたデータを利用した実行方法

台本生成、音声合成にはそれぞれにおよそ 2, 3 分の時間を要するため、展示などの際に、観客を待たせることになる。そのため、漫才の実行に必要なファイルの生成を予め済ませておき、そのデータを用いて漫才を実行できるようにした。

漫才のデータとして、生成した音声ファイル用のフォルダ、セリフファイル用フォルダ、台本ファイルを書き換えたテキストファイル、元になった XML ファイルをひとつのフォルダにまとめて保存する。このデータを再生するためのコンポーネントは前述のコンポーネントの一部を改良したもので、実行する漫才のデータを CUI で選択すると、漫才を実行することができる。

7 おわりに

本研究では、漫才ロボットを小型化するにあたって、RT ミドルウェア上で漫才を制御するシステムを開発した。漫才台本生成システムによって生成された台本をそのまま使用でき、小型化によって可搬性が向上し、稼動が容易となった。

本システムは、2014 年 8 月 19 日から 9 月 2 日までの大学都市 KOBE! 発信プロジェクト、2014 年 9 月 11, 12 日のイノベーション・ジャパン 2014 大学見本市などで展示を行い、甲南大学知能情報学部および漫才ロボットの紹介に貢献した。

今後は、現在 PC に接続して使用している小型漫才ロボットを Raspberry Pi などのマイコンを用いた制御ができるようにするなど、小型漫才ロボットの実用化に向けた改良を続けていく。

謝辞

本研究の一部は、平成 25 年度甲南大学父母の会の支援を得て、甲南大学学生ロボット工房の活動として行われた。また、本研究は私立大学等経常費補助金特別補助「大学間連携等による共同研究」の支援を受けた。甲南大学知能情報学部名誉教授 中山弘隆先生に感謝する。小型漫才ロボットのデザインはアートラボによって行われた。

参考文献

- [1] 宇野侑希, 田中聡, 伊藤洸, 宮崎光二, 北村達也, 永田亮, 灘本明代, 中山弘隆, “コミュニケーションロボット同士による漫才,” 甲南大学紀要知能情報学編, vol. 4, no. 2, pp. 167–180, 2011.
- [2] 真下遼, 梅谷智弘, 北村達也, 灘本明代, “つかみ・本ネタ・オチから構成される漫才ロボット台本自動生成手法の提案,” 第4回 Web インテリジェンスとインタラクション研究会 (ARG SIG-WI2) 論文集, pp. 59–64, 2014.
- [3] 真下遼, 灘本明代, “対立語抽出に基づく Web ニュースからの漫才ロボット台本自動生成手法の提案,” 第6回データ構造と情報マネジメントに関するフォーラム (DEIM Forum 2014), 8 pages, 2014.
- [4] Tomohiro Umetani, Ryo Mashimo, Akiyo Nadamoto, Tatsuya Kitamura and Hirotaka Nakayama, “Manzai robots: Entertainment robots based on auto-created manzai scripts from web news articles,” *Journal of Robotics and Mechatronics*, vol. 26, no. 5, pp. 662–664, 2014.
- [5] 梅谷智弘, 孝橋一希, 北村達也, 真下遼, 灘本明代, “分散駆動型ソフトウェアコンポーネントを利用した小型漫才ロボットの開発,” 日本機械学会ロボティクス・メカトロニクス講演会 2014 講演論文集, 3A1-T02, 3 pages, 2014.
- [6] 松本吉央, 山野辺夏樹 編, “特集：使える RT ミドルウェア,” 日本ロボット学会誌, vol. 28, no. 5, pp. 545–589, 2010.