

音声認識コンポーネント
(M211Listen)
ユーザマニュアル

目次

1. はじめに.....	4
2. 概要.....	4
3. 開発・動作環境.....	4
4. 制約事項.....	4
5. 音声認識コンポーネントの使用方法.....	5
5.1. コンポーネントの起動.....	5
5.2. コンポーネントの実行.....	6

図・表目次

図 1 ディレクトリ構成図(Windows).....	5
図 2 コンポーネント接続図.....	6
図 3 M211ListenTestOutComp 実行画面.....	7
表 1 開発・動作環境.....	4

- 改訂履歴 -

Ver.	更新日	改訂内容	改訂
1.0	2008/03/31	初版	イーガー
2.0	2010/07/26	OpenRTM1.0.0 対応	イーガー
3.0	2011/12/05	音声認識結果の信頼度の組み込み および発話意図認識の組み込み	ATR

1. はじめに

本書は、RT ミドルウェア上で動作する Windows 版 RT コンポーネント「音声認識コンポーネント」の使い方について示します。Linux 版に関しては、Ver.2 をご参照ください。

2. 概要

音声認識コンポーネント(以下本コンポーネント)は、入力された音声波形データを認識し、文字列で発話内容を出力します。Windows 版では、認識された文字列が感動詞「え」または「うん」の場合、これらが伝達する発話意図のカテゴリも出力します。

3. 開発・動作環境

音声認識コンポーネントの開発・動作環境を示します。

OS	WindowsXP SP3
開発ツール	Visual C++ 2008 Express Edition
使用ライブラリ	・OpenRTM-aist-1.0.0-RELEASE
ツール	・eclipse (RTSystemEditor にて使用) ・ATRASR

表 1 開発・動作環境

4. 制限事項

音声認識コンポーネントの制限事項を以下に示します。

- ATRASR 起動時に「エラー：メモリ不足」が表示された場合、ディレクトリ「ATRASR_V3/pref/」へ移動し、wav.ATRWecclrr の「ATRlattice:work_area=」設定値を低く設定してください。
- 音声認識結果の文字列は、文字コード『EUC-JP』形式です。
- 発話意図認識機能を使用した場合、音声認識が「え」または「うん」の感動詞を認識できなかった場合、発話意図認識結果は出力されません。

5. 音声認識コンポーネントの使用法

本章では、音声認識コンポーネントの使用法を、テストコンポーネントを使って説明します。

5.1. コンポーネントの起動

各コンポーネントの起動方法を説明します。なお、あらかじめ PC のマイク入力端子にマイクを接続しておいてください。

1) コンポーネントの起動準備

コンポーネントの起動は、Windows の場合には『図1 ディレクトリ構成図(WindowsXP)』の【任意のディレクトリ】で、行うようにしてください。

【任意のディレクトリ】	
└ 【ATRASR_V3】	
├ ATRASR_V3.lic	ライセンスキー
├ ATRfront.exe	音声認識モジュール(前処理)
├ ATRWclrr.exe	音声認識モジュール(認識エンジン)
├ GWPP.exe	音声認識モジュール(信頼度)
├ intermoduletool.exe	音声認識と発話意図認識モジュールの連携
├ paralingrec.exe	発話意図認識モジュール
└ 【model】	
├ 【am】	
├ └ HMNET.F.bin	男性用音響モデルファイル
├ └ HMNET.M.bin	女性用音響モデルファイル
├ └ └ gmm512_MMSE_16k_jp	雑音抑圧用モデル
├ └ └ gmmEPD.txt	発話区間検出用音声・雑音モデル
├ └ 【lm】	
├ └ └ test.bin	言語モデルファイル (FSA)
├ └ └ test.lex	辞書ファイル
└ 【pref】	
├ wav.ATRMSE	音声認識前処理(雑音除去)用 Config ファイル
├ wav.ATREpd	音声認識前処理(区間検出)用 Config ファイル
├ wav.ATRWclrr	音声認識エンジン用 Config ファイル
├ GWPP.bat	音声認識後処理(信頼度)用 Config バッチファイル
├ config_socket_epd.txt	音声認識と発話意図認識の連携用 Config ファイル
├ config_socket_epd_paraling.txt	音声認識と発話意図認識の連携用 Config
├ └ config_paraling.txt	発話意図認識接続用 Config ファイル
├ M211ListenComp.exe	音声認識コンポーネント
├ InMicWinComp.exe	音声入力コンポーネント
├ M221TimeServerComp.exe	時間管理コンポーネント
├ M211ListenTestOutComp.exe	音声認識結果出力コンポーネント
└ rtc.conf	コンフィグファイル

図 1 ディレクトリ構成図(Windows)

※モジュール起動ディレクトリから【ATRASR_V3】ディレクトリ内のファイルを相対パスにて参照しています。モジュール起動ディレクトリから【ATRASR_V3】ディレクトリ、及び配下のサブ・ディレクトリ、ファイルへの相対パスが変わらないようにしてください。

※音声認識ソフトウェア「ATRASR_V3」については、次世代ロボット知能化プロジェクト内において、非商用利用に限り無償での提供が可能ですが、別途 秘密保持契約 を結んでいただく必要があります。ATRASR_V3 には、ライセンスキー ATRASR_V3.lic、実行ファイル ATRfront.exe および ATRWcclrr.exe、音響モデル HMMNet.F.bin および HMMNet.M.bin が含まれます。これらのファイルを図1に示された正しいパスにコピーしてください。

2) 接続・設定

RTSystemEditorを起動し、各コンポーネントを次の図のように接続します。発話意図認識も起動させる場合は、M211Listen のパラメータ設定で、“paraling”を “1” に設定してください。音声認識のみを起動させる場合は、“paraling”を “0” にしてください。

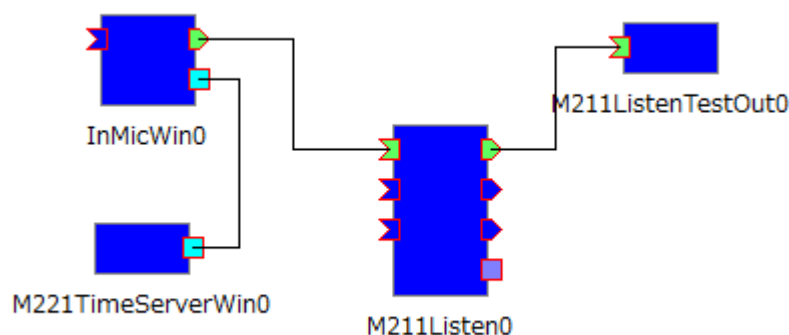


図 2 コンポーネント接続図

3) 起動

RTSystemEditor の右クリックメニューから All Activate を選択します。

5.2. コンポーネントの動作確認

- 1) 音声認識結果出力コンポーネント『M211ListenTestOutComp』を起動しているターミナルに、『図 3 M211ListenTestOutComp 実行画面』のような認識結果が出力されます。発話ごとに音声認識結果の文字列、次の行に対応する認識結果の信頼度 (0~1 の値: に近い程信頼度が高い) が表示されます。発話意図認識機能を起動している場合は、「えー」および「うん」の感動詞が認識された場合、これらの発話意図のカテゴリ結果が次の行に表示されます。



図 3 M211ListenTestOutComp 実行画面

以上