

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.....	7

図・表目次

図 1 ディレクトリ構成図(Windows).....	5
図 2 ディレクトリ構成図(Ubuntu).....	6
図 3 コンポーネント接続図.....	7
図 4 InListenStateComp 実行画面.....	8
図 5 M211ListenTestOutComp 実行画面.....	8
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

1. はじめに

本書は、RTミドルウェア上で動作する 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。

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。

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	Ubuntu 10.04	WindowsXP SP3
開発ツール	gcc 4.4.3 (Ubuntu 4.4.3-4ubuntu5)	Visual C++ 2008 Express Edition
使用ライブラリ	・OpenRTM-aist-1.0.0-RELEASE ・libstdc++.so.5	・OpenRTM-aist-1.0.0-RELEASE
ツール	・eclipse (RTSystemEditor にて使用) ・ATRASR	・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/pref/」へ移動し、config.ATRFwcclrr、wav.ATRFwcclrr の「ATRlattice:work_area=」設定値を低く設定してください。
- Windows 環境において本コンポーネントを使用する場合は、付属のコンフィグファイル「configS.ATRFwcclrr」、「wavS.ATRFwcclrr」を ATRASR の prefs ディレクトリに置いてください。
※ p.5『図1 ディレクトリ構成図(Windows)』参照
- 音声認識結果の文字列は、文字コード『EUC-JP』形式です。

5. 音声認識コンポーネントの使用方法

本章では、音声認識コンポーネントの使用方法を、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、Ubuntu の場合には『図2 ディレクトリ構成図(Ubuntu)』の【M211Listen】で、それぞれ行うようにしてください。

【任意のディレクトリ】	
└ 【ATRASR】	
├ └ ATRFwcclrr	音声認識モジュール
├ └ ATRASR_v3.lic	ライセンスキー
├ └ 【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	雑音抑圧用クリーン音声モデル
├ │ └ 【lm】	
├ │ │ └ Robovie.bin	言語モデルファイル (FSA)
├ │ │ └ Robovie.lex	辞書ファイル
├ └ 【pref】	
├ │ └ config.ATRFwcclrr	Raw データ用 Config ファイル
├ │ └ configS.ATRFwcclrr	Raw データ用 Config ファイル
├ │ └ wav.ATRFwcclrr	Wav データ用 Config ファイル
├ │ └ wavS.ATRFwcclrr	Wav データ用 Config ファイル
└ M211ListenComp.exe	音声認識コンポーネント
└ InMicComp.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
└ InListenStateComp.exe	音声認識状態設定コンポーネント
└ 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)

```

【任意のディレクトリ】
└─ 【ATRASR】
   │   └─ ATRFwcclrr          音声認識モジュール
   │   └─ ATRASR_v3.lic      ライセンスキー
   │   └─ 【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 雑音抑圧用クリーン音声モデル
   │       │       └─ 【lm】
   │       │           │   └─ Robovie.bin    言語モデルファイル (FSA)
   │       │           │       └─ Robovie.lex 辞書ファイル
   │       └─ 【pref】
   │           │   └─ config.ATRFwcclrr      Raw データ用 Config ファイル
   │           │       └─ wav.ATRFwcclrr      Wav データ用 Config ファイル
   └─ 【M211Listen】
       │   └─ M211ListenComp      音声認識コンポーネント
       │   └─ InMicComp           音声入力コンポーネント
       │   └─ M221TimeServerComp   時間管理コンポーネント
       │   └─ InListenStateComp    音声認識状態設定コンポーネント
       │   └─ M211ListenTestOutComp 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 2 ディレクトリ構成図(Ubuntu)

※モジュール起動ディレクトリから【ATRASR】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】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。

2) 接続

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。

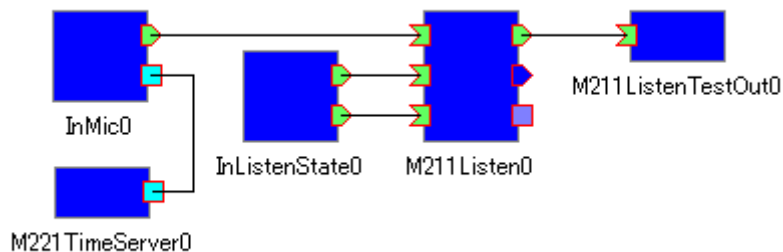


図 3 コンポーネント接続図

3) 起動・設定

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

<Ubuntu10.04 の場合>

音声認識結果出力コンポーネント『M211ListenTestOutComp』を起動したターミナルで、ターミナルの文字コードの設定を日本語 (EUC-JP) に設定します。

※P.4 の制限事項を参照

5.2. コンポーネントの実行

各コンポーネントの実行・操作方法を説明します。

- 1) 音声認識状態設定コンポーネント『InListenStateComp』を起動しているターミナルに、『図 4 InListenStateComp 実行画面』のような入力画面が表示されます。ここでまず “1” と入力すると、音声認識が開始されます。次に、マイクに対して発話した後、“0” を入力して音声認識を終了します。

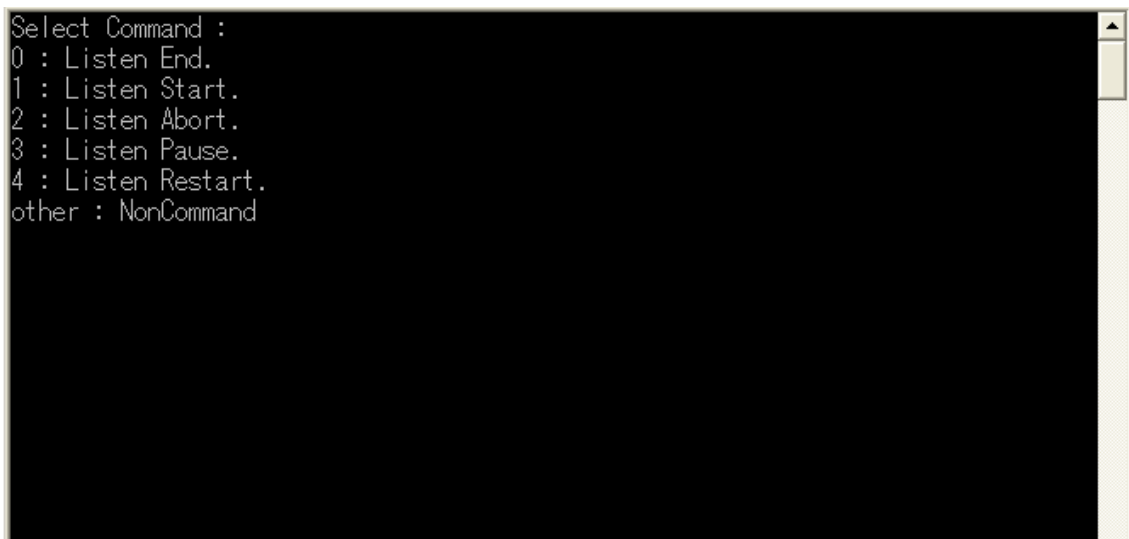


図 4 InListenStateComp 実行画面

- 2) 認識が完了すると、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 5 M211ListenTestOutComp 実行画面』のような認識結果が出力されます。



図 5 M211ListenTestOutComp 実行画面

以上