Harebot机器人平台1.0

开发手册

之语音模块

**中国科学院软件所**

目录

前言

欢迎使用Harebot机器人平台（Harebot Robot Platform，HRP）！

Harebot机器人平台是基于中科院软件所自主研发的Harebot机器人，开发的一款符合XXX使用的应用开发平台，提供XXX、XXX等服务，为智能机器人应用开发爱好者提供方便易用的开发平台，使得用户能够基于该平台进行多种智能应用开发，平台主要特点有：

1. XXX

……

本手册主要讲述如何利用Harebot机器人平台……，其适用的读者有：

……

本手册基本内容：

1. 功能概述
2. 代码分析

……

1. 功能概述

语音模块用于获取用户语音指令，给机器人装上耳朵。实现类人“听力”。模块集成科大讯飞“命令词识别”“语音合成”两组语音服务开发包，实现识别用户指令，并以语音形式将识别结果以及执行动作反馈给用户。

（一）模块具体功能如下：

1. 录制用户语音指令，并保存成.wav格式音频。
2. 对用户指令音频进行内容识别。
3. 将识别结果映射成相应的执行动作，并将其对外发布出去。
4. 相关client订阅发布出来的指令，并执行相应的动作；在当前版本中，client接收到发布的执行后，实现将动作内容以语音形式反馈给用户。
5. 相关说明：
6. 音频录制格式：

采样率16k/8k；

采用精度16bit；

单声道；

wav或者pcm格式；

1. 命令词识别语法可定制：

离线识别的命令词是开发者自己定义，命令词最大长度为31个字，个数为215个。需要先构建语法，然后指定使用的语法。

其中，简单的语法示例如下：

<commands>:(找一下|打电话给) <name>;

<name>: 张三|李四;

. . . . . . .

该语法使识别引擎可以支持以下说法：找一下张三 、打电话给张三 、找一下李四 、打电话给李四。凡是用户说出这个范围中的任意一句话，均可以被识别系统识别。如果用户说的话不在上述范围中，识别系统可能拒绝识别。

1. 代码分析

本模块代码主要包含以下：

1. user\_command\_client.cpp
2. user\_command\_server.cpp
3. command\_subscriber.cpp

1、user\_command\_server.cpp

在该文件中，建立“user\_command\_server”节点，提供“UserCommand.srv“服务，具体如下：

1）启动录音，并将录制的音频保存成.wav文件；

2）对音频内容进行识别；

3）发布识别的结果；

这里重点说明关于语音识别部分的代码。在该部分代码中需要调用一些通用接口进行基本的登陆登出、数据传输；自动语音识别（Automatic Speech Recognition,ASR）接口，用于开始一路会话，根据传入的语法，进行语音识别，获取识别结果，结束该路回话。

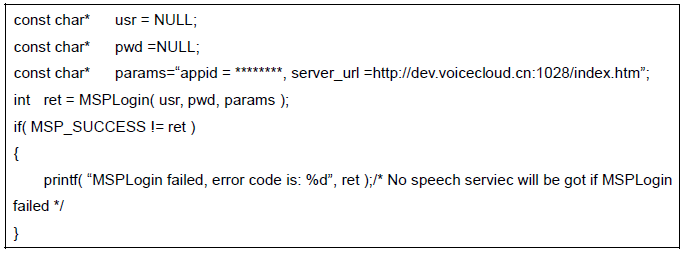
* 1. 通用接口：

关于通用接口，提供一下函数调用：

其返回值说明如下：

对于开发接口，如果调用成功，返回值为int 型的接口都会返回MSP\_SUCCESS，否则返回错误代码，错误代码参见msp\_errors.h；返回值为指针类型的接口，函数执行失败返回空指针0，错误代码可以通过相关的回传参数查看，函数执行成功但无数据时返回空指针0，有数据时返回非0 指针。

用法示例：



1.2 ASR接口函数列表

关于ASR提供如下函数调用：

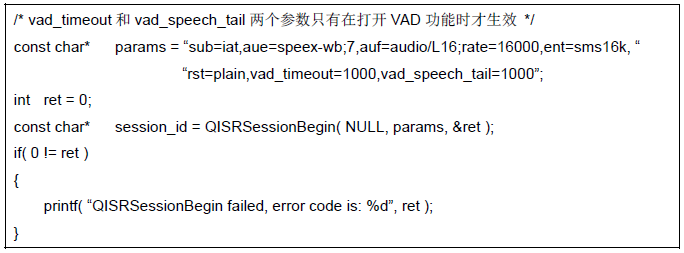


其返回值说明：

对于开发接口，如果调用成功，返回值为int 型的接口都会返

回MSP\_SUCCESS，否则返回错误代码，错误代码参见msp\_errors.h；返回值为指针类型的接口，函数执行失败返回空指针0，错误代码可以通过相关的回传参数查看，函数执行成功但无数据时返回空指针0，有数据时返回非0 指针。

用法示例：



另，“user\_command\_server”节点在成功进行语音识别之后，将识别结果通过“Command.msg“发布给各个订阅者。

2、user\_command\_client.cpp

在该文件中，建立“user\_command\_client”节点，通过调用

“UserCommand.srv “服务触发“user\_command\_server”节点，提供相应的服务内容。

3、command\_subscriber.cpp

在该文件中，建立“subscriber “节点，用于订阅由“user\_command\_client”节点发布的“Command.msg “消息。在当前版本中，我们加入了一个占位功能，即：接收到识别结果后，执行相应的内容动作，这里，我们将识别到的语音文本内容，再次合成为音频文件，并播放给用户。

因此，这里涉及到了语音合成（Text To Speech，tts）接口。

3.1 TTS接口说明

关于TTS提供如下函数调用：

其返回值说明：

对于开发接口，如果调用成功，返回值为int 型的接口都会返回MSP\_SUCCESS，否则返回错误代码，错误代码参见msp\_errors.h；返回值为指针类型的接口，函数执行失败返回空指针0，错误代码可以通过相关的回传参数查看，函数执行成功但无数据时返回空指针0，有数据时返回非0 指针。

用法示例：

