

Rokid语音识别

版本: s2

1. 概述

Rokid语音识别系统由以下几部分组成。

1. 拾音及识别服务 (BearKid)

2. tts播放服务

由'语义处理服务'根据需要调用。

3. 媒体播放服务

由'语义处理服务'根据需要调用。比如播放特定的音乐url

4. 语音信息及识别结果处理 (SpeechExecutor)

获取识别结果，发送给语义结果处理服务 获取声音方向信息，根据不同设备做出不同的用户反馈 获取音强信息，根据不同设备做出不同的用户反馈 获取异常信息，根据不同设备做出不同的用户反馈

5. 语义结果处理服务 (CloudAppClient)

根据语义结果(nlp, action)，执行相应的操作，如播放tts语音，播放音乐，播放新闻等

注：SpeechExecutor, CloudAppClient根据不同的用户需求做定制，在这里提供的这两个模块的源码可以看作是demo性质。

BearKid, tts, 媒体播放是稳定的服务提供者，在本文档接下来的部分会说明接口定义。

- Rokid语音识别系统会分为远场拾音及近场拾音两个版本，本源码编译出的是近场拾音版本。
- 远场及近场接口定义一致，某些接口仅在近场拾音时有用，某些接口仅在远场拾音中有用，在下面的接口定义中会标明。

2. 源码使用说明

1. 获取源码及其子模块全部内容

```
git clone <源码git url> -b dev
cd <源码根目录>
./initsub.sh
```

- 2. 打开eclipse (需安装好android开发套件)
- 3. 选择菜单File-->Import
- 4. 在弹出的对话框中选择Android-->Existing Android Code Into Workspace
- 5. 在弹出的对话框中选定我们的源码目录， 点击Finish按钮

3. 接口定义

3.1 BearKid (Speech+Siren)

- control

发出开关拾音等指令

~	名称	类型	描述
参数	action	int	1: 激活(idle-->active)(远场) 2: 休眠 (active-->idle)(远场) 3: 静音 (任意-->mute)(远场) 4: 取消静音 (mute-->idle)(远场) 5: 开始拾音(近场) 6: 停止拾音(近场)

- poll

获取结果。结果分为以下类型：

- 声音方向信息(远场)
- 激活确认 (远场)
- 中间asr (近场、远场)
- 结果asr (近场、远场)
- 语义结果 (近场、远场)
- 退出激活 (远场)

声音信息 (声强) (远场)

远场专有异常

未连网激活, cancel

近场、远场共有异常

timeout, 服务错误, 其它错误待定

~	名称	类型	描述
返回值		BearKidResult	语音服务结果

• BearKidResult数据结构

~	名称	类型	描述
	type	int	类型 0: 寻向信息 1: 语音活动信息 2: 激活确认 3: 中间asr 4: 结束asr 5: 语义结果 6: 退出激活 7: 异常
	location	double	声音角度
	activation	int	0: accept 1: reject
	asr	string	
	nlp	string	
	action	string	
	energy	double	音强
	extype	int	异常类型: 激活但未连网 语音识别取消 本地超时 服务错误

3.2 Tts

- speak

文字转语音请求

~	名称	类型	描述
参数	text	string	文字
参数	callback	TtsCallback	回调接口
返回值		int	speak请求id

- cancel

取消文字转语音请求

~	名称	类型	描述
参数	id	int	speak请求id

- cancelByOwner

取消当前进程所有文字转语音请求

- pause

暂停播放语音

- resume

继续播放语音

3.3 MediaPlayer

3.3.1 播放器调用接口（IRKMediaPlayer）

- void setVideoPath(in String url);

设置播放资源

~	名称	类型	描述
参数	url	String	播放地址

- void start();

开始播放

- void seekTo(in int position);

跳转到指定的播放位置

~	名称	类型	描述
参数	position	int	播放位置(单位：毫秒)

- void pause();

暂停播放

- void stop();

停止播放

- boolean isPlaying();

是否处于播放状态

- boolean canPause();

是否可以暂停

- int getDuration();

获取总时长

- int getPosition();

获取当前播放位置

- void setMediaStateCallback(in MediaStateCallback mediaStateCallback);

设置播放状态回调接口

~	名称	类型	描述
参数	mediaStateCallback	MediaStateCallback	播放器状态回调接口

3.3.2 播放器状态回调接口（MediaStateCallback）

- void onStartPlay();

播放器开始播放

- void onTruckTimeout();

播放器播放过程卡顿（5秒）超时

- void onPause(int position);

播放器暂停

~	名称	类型	描述
参数	position	int	播放暂停位置(单位：毫秒)

- void onStop();

播放器停止播放

- void onCompletion();

播放器播放完成

- void onError(int what, int extra);

播放器播放出错

~	名称	类型	描述
参数	what	int	播放出错类型
参数	extra	int	播放出错码

3.4 BearKidAdapter

帮助访问BearKid服务的工具类

3.4.1 调用接口 (BearKidAdapter)

- initialize

初始化

~	名称	类型	描述
参数	context	Context	android应用context
参数	callback	BearKidCallback	回调接口
返回值		boolean	

- close

关闭

- control

调用BearKid服务control方法

3.4.2 回调接口 (BearKidCallback)

- onVoiceEvent

寻向信息

语音活动信息

激活确认/退出激活

~	名称	类型	描述
参数	event	int	事件类型: 1: EVENT_LOCATION 语音方向 2: EVENT_VOICE_INFO 音强信息 3: EVENT_ACTIVATION 语音激活结果 4: EVENT_DEACTIVE 退出激活状态
参数	arg1	int	'event' == EVENT_ACTIVATION, 'arg1'表示激活结果。 ACTIVATION_ACCEPT ACTIVATION_REJECT
参数	arg2	double	'event' == EVENT_LOCATION, 'arg2'表示语音方向 'event' == EVENT_VOICE_INFO, 'arg2'表示音强

- onIntermediateResult

中间asr

~	名称	类型	描述
参数	asr	string	

- onRecognizeResult

asr, nlp, action

~	名称	类型	描述
参数	asr	string	整句语音转文字结果 如果'asr'不为空，nlp action一定为空
参数	nlp	string	语义理解结果 'nlp' 'action'同时为空或不为空 如果'nlp'不为空，'asr'一定为空
参数	action	string	rokid cloud app专用语义结果

- onException

~	名称	类型	描述
参数	exception	int	异常类型 1: EXCEPTION_ACTIVATE_NO_INET 2: EXCEPTION_CANCEL 3: EXCEPTION_TIMEOUT 4: EXCEPTION_SERVICE_ERROR