

Rokid语音识别

版本: s3

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，根据业务需求定制行为

获取音强信息，根据业务需求定制行为

获取异常信息，根据业务需求定制行为

注: 源码中的SpeechExecutor是demo性质，没有做任何业务需求定制

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. 源码获取及编译

2.1 准备编译环境

- 安装openjdk

安装方法查看openjdk的官方说明。需要1.7以上版本。(为后继步骤的方便，最好装1.8以上的版本)

mac上无方便的安装openjdk方法，可使用oracle jdk 1.8的版本。

- 安装android sdk命令行环境

打开[android开发官方网站](#)

拉到页面最下方“仅获取命令行工具”，找到对应你操作系统的工具包下载并解压。

注：我们源码只支持mac及linux平台编译

下面以<android_sdk_dir>指代刚刚下载的android sdk命令行工具目录。

```
cd <android_sdk_dir>/tools/bin
./sdkmanager "ndk-bundle"           安装ndk命令行工具
./sdkmanager "build-tools;27.0.1"   安装build-tools 最新的27.0.1版本
./sdkmanager "platforms;android-23" 获取android6.0的android.jar
```

注：sdkmanager的运行需要openjdk1.8环境，但我们源码的编译只需要openjdk1.7以上的环境。

如果sdkmanager运行出错，检查你的JAVA_HOME环境变量，看是否指向jdk1.8以上的版本。

sdkmanager运行完以后，openjdk1.8的环境不是必须的

如果在编译本源码时，使用的是openjdk1.8环境，那么安装的build-tools版本需要27以上，如果使用的是openjdk1.7环境，build-tools版本只需在23以上即可

android.jar只需android6.0以上的版本，可自行选择

以上步骤执行成功后，下会出现platforms, build-tools, ndk-bundle子目录。

配置环境变量

```
export ANDROID_SDK_ROOT=<android_sdk_dir>
export ANDROID_BUILD_TOOL_VERSION=27.0.1    根据你安装的build-tools版本指定
```

2.2 获取源码

```
mkdir <vsys_src>
cd <vsys_src>
repo init -u ssh://<user>@s.rokid-inc.com:29418/open-platform/client/manifests/vsy
s --no-repo-verify
repo sync
```

现在还是公司内部仓库，未来将开源到github

2.3 编译

```
cd <vsys_src>
. build/envsetup.sh
lunch
    lunch将列出三个选项如下：
    1. android6.0/7.0/7.1
    2. android4.4
    3. android5.0/5.1
    根据需要自行选择
make
```

编译后apk输出在 out/target/<target_platform>/nearfield/RKVoiceSystem.apk

现在只有近场拾音版，远场拾音版后继加入

<target_platform>为19/21/23。19对应android4.4， 21对应android5.0/5.1， 23对应android6.0/7.0/7.1

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

- addCustomWord

增加自定义激活词

~	名称	类型	描述
参数	type	int	1: 激活词 2: 退出词
参数	word	String	激活词汉字(utf8编码)
参数	pinyin	String	激活词拼音
返回值		int	0: 成功 < 0: 错误码

- removeCustomWord

删除自定义激活词

~	名称	类型	描述
参数	word	String	
返回值		int	0: 成功 < 0: 错误码

- queryCustomWord

获取提定类型的自定义激活词

~	名称	类型	描述
参数	type	int	1: 激活词 2: 退出词
返回值		List	所有自定义激活词列表

- 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

暂缺，待补完

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

- addCustomWord

增加自定义激活词

~	名称	类型	描述
参数	type	int	1: 激活词 2: 退出词
参数	word	String	激活词汉字(utf8编码)
参数	pinyin	String	激活词拼音
返回值		int	0: 成功 < 0: 错误码

- removeCustomWord

删除自定义激活词

~	名称	类型	描述
参数	word	String	
返回值		int	0: 成功 < 0: 错误码

- queryCustomWord

获取指定类型的自定义激活词

~	名称	类型	描述
参数	type	int	1: 激活词 2: 退出词
返回值		List	所有自定义激活词列表

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