

京东 AI 语音识别 Android SDK

用户接入文档

目录

1	简介.....	3
1.1	基础功能	3
2	使用说明.....	3
2.1	开发前准备.....	3
2.2	支持的设备.....	3
3	集成指南.....	3
3.1	各权限使用说明	4
4	SDK 接口定义	4
4.1	语音识别对象(SpeechManager)	4
4.1.1	创建对象	4
4.1.2	设置回调监听	5
4.1.3	发送识别请求	5
4.2	语音识别回调接口(SpeechListener).....	7
4.2.1	回调方法	7
5	示例 demo	7
6	常见问题及错误码描述	9

1 简介

京东语音 Android SDK 是集成京东 AI 语音相关基础服务，运行在 Android 设备上的一体化开放工具包。

1.1 基础功能

当前主要支持的功能如下：

录音，语音数据处理，语音端点检测，网络通讯，状态通知，实时反馈语音能量，识别结果返回等。

2 使用说明

2.1 开发前准备

在正式开发之前，需要为应用申请产品 ID，请发邮件到：jdai-speech@jd.com。

如果在接入开发过程中有其他问题，可以联系我们的技术支持：asr-support@jd.com。

2.2 支持的设备

系统: 支持 Android 4.3 及以上系统。通过 minSdkVersion 参数检测。

机型: 手机、平板及其它智能终端设备。

架构: 支持 armeabi, armeabi-v7a, armeabi-v8a, x86, x86_64。

外设: 设备上有麦克风输入。

网络: 支持 wifi 及系统网络。

3 集成指南

- 1) 将 jar 包拷贝到工程 libs 目录下，并加入工程依赖。
- 2) 将 libs 目录下 armeabi, arm64-v8a, armeabi-v7a, x86 文件夹按需添加到 android studio 工程的 src/main/jniLibs 目录中, eclipse 工程默认为 libs 目录下面。
- 3) 为工程添加必要的权限

```
<uses-permission android:name="android.permission.INTERNET" />
<uses-permission android:name="android.permission.RECORD_AUDIO" />
<uses-permission android:name="android.permission.ACCESS_NETWORK_STATE" />
<uses-permission android:name="android.permission.READ_PHONE_STATE" />
```

3.1 各权限使用说明

名称	用途
android.permission.INTERNET	允许应用联网发送语音数据给服务器获取识别结果
android.permission.RECORD_AUDIO	运行使用麦克风采集录音
android.permission.ACCESS_NETWORK_STATE	获取当前 WiFi 状态
android.permission.READ_PHONE_STATE	获取 imei 作为唯一标识

Android 6.0 及以上的动态申请权限,示例参考.

```
private void applyPremission() {
    String permissions[] = {
        Manifest.permission.RECORD_AUDIO,
        Manifest.permission.ACCESS_NETWORK_STATE,
        Manifest.permission.INTERNET,
        Manifest.permission.READ_PHONE_STATE };
    ArrayList<String> applyList = new ArrayList<String>();
    for (String perm : permissions) {
        if (PackageManager.PERMISSION_GRANTED !=
            ContextCompat.checkSelfPermission(this, perm)) {
            applyList.add(perm);
        }
    }
    String tmpList[] = new String[applyList.size()];
    if (!applyList.isEmpty()) {
        ActivityCompat.requestPermissions(this,
            applyList.toArray(tmpList), 123);
    }
}
```

4 SDK 接口定义

4.1 语音识别类(SpeechManager)

4.1.1创建对象

SpeechUtility.create(Context context, EngineType engineType)

名称	描述
context	上下文对象
engineType	引擎类型，ASROperator.START 语音识别

4.1.2 设置回调监听

setListener(SpeechListener listener)

参数说明：

名称	描述
listener	语音识别回调接口，通知外部获取识别结果状态

4.1.3 发送识别命令

send(String cmd, String params)

参数说明：

名称	描述	取值
cmd	语音识别命令	ASR.START 开始识别 ASR.CANCEL 取消识别 ASR.STOP 停止识别
params	识别设置的参数，JSON 格式的字符串	各个参数设置见下方(识别 JSON 参数说明)

识别 JSON 参数说明：

JSON 字段	描述	类型	取值	必填
SAMPLE	录音采样率	int	目前只支持 16000	否
DOMAIN	领域模型设置	String	search (默认)	否
SERVER_VAD_ENABLE	是否开启云端 VAD	boolean	true 开启 false 不开启 (默认)	否
AUDIO_SOURCE	语音识别的录音来源	String	音频文件的字符串路径 "/sdcard/Audio.pcm"，不设置从麦克风采集录音 (默	否

			认)	
APPID	产品 ID	String	产品 ID 请发邮件到 jdai-speech@jd.com 申请	是
LOCAL_VAD_ENABLE	是否启用离线 vad	boolean	true 启用 false 不启用(默认)	否
POST_PROCESS	是否启用数字后处理	int	1 启用(开启后会将结果中的数字汉字转换成阿拉伯是数字(例如: “一千”会转换成”1000”) 0 不启用(默认)	否
VAD_START_DELAY	(LOCAL_VAD_ENABLE 为 true 的情况下生效)离线 vad 前端点超时时间(单位:秒)	double	[0.25~0.5] 默认 0.25	否
VAD_END_DELAY	(LOCAL_VAD_ENABLE 为 true 的情况下生效)离线 vad 后端点超时时间(单位:秒)	double	[0.4 ~ 2.0] 默认 0.5	否
LONG_SPEECH	是否开启长语音识别, 开启长语音之后会持续一直识别, 调用 ASR.STOP 关闭本次识别	boolean	true 开启 false 不开启(默认)	否
PHRASE_LIST	上传个性化词列表	JSONArray	上传的词表总长度小于 2K, 上传示例:[“王茗”, “班玛县”, “龙羊峡”], 在用户说“打电话给王茗”, “导航到班玛县”, “我要去龙羊峡”会优先识别出个性化词表里面的词, 默认为空。	否
PARTIAL_RESULT	是否开启实时结果返回	int	1 启用 0 不启用(默认)	否

4.2 语音识别回调接口(SpeechListener)

4.2.1回调方法

public void onEvent(SpeechEvent event, String result, byte[] data)

参数说明

名称	描 述	取 值
event	回调事件名称	SPEECH_START: 准备就绪,开始录音 SPEECH_BEGIN 用户开始说话 SPEECH_END 用户结束说话 SPEECH_VOLUME 实时音量大小回调 SPEECH_PARITAL_RESULT 临时识别结果 SPEECH_RESULT 最终结果
result	相关回调事件返回的json字符串结果	具体见下方"result 参数 JSON 字符串说明"
data	相关事件返回携带的音频数据，例如保存录音音频到本地	

result 参数 JSON 字符串说明

(1) SPEECH_START、SPEECH_BEGIN 与 SPEECH_END 事件时没有 result 返回。

(2) SPEECH_VOLUME 事件返回 JSON 字符串结果说明与示例

名称	类型	描述
volume	String	音量大小,取值范围[0~100]

示例：

```
{"volume": "20"}
```

(3) SPEECH_PARITAL_RESULT 与 SPEECH_RESULT 事件返回 JSON 字符串结果说明与示例

名称	类型	描 述
err_code	int	错误码
request_id	String	请求 ID

err_msg	String	错误信息
context	JSON	识别结果

示例：

```
{
  "err_code":"0",
  "request_id":"fe0c804d-dec9-4a80-8167-b730b697ee5b",
  "err_msg":"识别成功",
  "context":[{"text": "今天天气怎么样"}]
}
```

5 示例 demo

```
SpeechManager asrManager = SpeechUtility.create(this, EngineType.ASR);
JSONObject intent = new JSONObject();
//设置采样率
intent.put(SpeechConstant.SAMPLE_RATE, 16000);
//是否启用云端 vad
intent.put(SpeechConstant.SERVER_VAD_ENABLE, true);
//设置 APPID, 请邮件到 jdai-speech@jd.com 申请
intent.put(SpeechConstant.APPID, "1234567890");
//设置回调监听
asrManager.setListener(new SpeechListener() {
    @Override
    public void onEvent(SpeechEvent event, String result, byte[] data) {
        switch (event) {
            case SPEECH_START:
                //准备就绪, 开始录音
                break;
            case SPEECH_BEGIN:
                //用户开始说话
                break;
            case SPEECH_END:
                //用户结束说话
                break;
            case SPEECH_VOLUME:
                //用户说话的音量返回值
                break;
            case SPEECH_PARTIAL_RESULT:
                //中间识别结果
                break;
        }
    }
});
```



```
        case SPEECH_RESULT:  
            //最终识别结果  
            break;  
    }  
}  
});  
//启动识别  
asrManager.send(ASROperator.START, intent.toString(), null, 0, 0);
```

6 常见问题及错误码描述

err_code	err_msg	描述
0	识别成功	~
-1001	录音异常	录音机被占用或者录音设备异常
-1002	权限异常	未声明录音权限或者权限被其他软件限制
-1003	网络连接失败	没有网络权限或者网络未连接
-1004	网络超时	网络质量差
-1005	语音太长	说话时间超出限制,目前只支持 60s 的语音录入
-1006	没有识别到语音	用户没有说话
-1007	服务器出错	服务器识别过程出错
-1008	客户端错误	参数配置出错,集成过程中参数设置出错,请参照开发文档检查
-1009	语音格式错误	SDK 只接收 pcm 或者 wav 格式的语音
-1010	Appld 鉴权失败	Appld 没有设置或者设置的不合法.请邮件到 jdai-speech@jd.com 进行申请
-1011	语音文件头出错	语音文件头信息校验出错, 请检查语音是否正常
-1012	语音采样率或者通道错误	SDK 要求语音采样率 16000 单通道录音.
-1013	vad 初始化失败	vad 资源配置错误
-1014	新网关鉴权出错	Appkey 或者 SecretKey 填写错误