

京东AI语音识别iOS_SDK

用户接入手册

目录

1. 概述	3
1.1 基础功能	3
2. 兼容性	3
3. 资源占用描述	3
4. 集成	4
4.1 Demo 工程	4
4.2 开发前准备	4
4.3 获取麦克风权限	4
4.4 添加头文件和静态库	4
5. 语音 SDK 接口说明	4
6. 常见问题及错误码描述	8

1. 概述

本文档是京东语音 iOS SDK 的用户指南，SDK 集成了京东 AI 语音相关基础服务，描述了语音识别等相关接口的使用说明。

1.1 基础功能

当前主要支持的功能：

基本功能：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. 兼容性

类别	兼容范围
系统	支持 iOS8.0 及以上系统
架构	armv7、arm64、i386、x86_64
网络	支持移动网络、WIFI 等网络环境
外设	设备有麦克风输入
开发环境	建议使用最新版本 Xcode 进行开发

3. 资源占用描述

SDK 类型	文件
语音识别 JD_Speech.framework	JDAISASRDefines.h
	JDAISASRParameters.h
	JDAISSpeechManager.h
	libJD_Speech
VAD 模型文件	JD_Speech_vad.bin

开启和关闭 BITCODE 对于最终编译大小影响不大，编译开启了 BITCODE 选项。

由于 BITCODE 开启会导致二进制文件体积增大，这部分会在 AppStore 发布时进行进一步编译优化，并不会引起最终文件的体积变化。

虽然 ipa 的包会有增大，但实际在 AppStore 的优化下，用户下载的应用大小会远小于 ipa 文件的大小。

4. 集成

4.1 Demo 工程

强烈建议在使用 iOS SDK 之前，运行并试用 Demo 工程相关功能，参考 Demo 工程的调用和配置的方式。

1. 双击使用 Xcode 打开 JDASR_Speech_Test/JD_Speech_Test.xcodeproj
2. 确保网络通畅，运行工程

4.2 开发前准备

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

4.3 获取麦克风权限

首先要为 APP 声明使用麦克风权限，在新建的应用下面找到 info.plist，然后在其中新增麦克风权限的声名：

```
<key>NSMicrophoneUsageDescription</key>
<string>microphoneDescription</string>
```

这样添加了之后，系统就知道了 APP 要访问系统的麦克风权限。

4.4 添加头文件和静态库

JD_Speech_vad.bin 模型文件，将拖入工程目录作为 Copy Bundle Resources。

JD_Speech.framework 提供的是静态库，开发者只需要将库文件拖入工程目录即可，可参考 demo 程序的引用。

5. 语音 SDK 接口说明

详细接口代码请参考 SDK 头文件：

```
JDAISSpeechManager.h
JDAISASRDefines.h
JDAISASRParameters.h
```

语音识别类：

```
JDAISSpeechManager
```

初始化语音识别对象：

```
// engineType
// 引擎类型，语音识别为 JDAIS_ENGINETYPE_ASR
+ (instancetype) create : (NSString *)engineType;
```

获取版本号：

```
- (NSString *)getSDKVersion;
```

设置监听回调：

```
// setListener
// 语音状态回调接口
// 请参考 JDAISASRDefines.h 中的 JDAISClientASRDelegate 协议并且实现相关接口。
- (void) setListener : (id)setListener;
```

发送识别请求：

```
// cmd
// 语音识别命令，请参考 JDAISASRDefines.h 中定义。
// 开始识别： JDAIS_ASR_CMD_START
// 取消识别： JDAIS_ASR_CMD_CANCEL
// 停止识别： JDAIS_ASR_CMD_STOP
// params
// 参数列表，JSON 格式类型，具体参数定义参考 JDAISASRParameters.h 中定义。
- (void) send : (NSString *)cmd withParams:(NSString *)params;
```

params 参数列表字段说明

名称	类型	必填	说明
JDAIS_ASR_SAMPLE	NSInteger	否	录音采样率，8000/16000，默认为 16000
JDAIS_ASR_DOMAIN	NSString	否	领域模型，默认为"search"
JDAIS_ASR_SERVER_VAD_ENABLE	bool	否	是否开启云端 VAD 截断，0 为关闭，1 为开启，默认为 0 关闭
JDAIS_ASR_APPID	NSString	是	APPID 鉴权相关，发送邮件到 jdai-speech@jd.com 进行申请
JDAIS_ASR_APPKEY	NSString	是	权限认证使用
JDAIS_ASR_SECRETKEY	NSString	是	权限认证使用
JDAIS_ASR_AUDIO_SOURCE	NSString	否	语音识别的音频来源，不设置将从录音机输入，设置成文件路径格式："/sdcard/Audio.pcm"
JDAIS_ASR_RECOGNIZE_TYPE	NSInteger	否	实时语音识别类型，0 为实时识别，1 为一句话识别，默认为 1
JDAIS_ASR_IS_NEED_RECORD_FILE	NSInteger	否	是否需要保存录音文件，0 为不保存，1 为保存，默认为 0 并且不进行录音文件保存

JDAIS_ASR_RECORD_FILE_PATH	NSString	否	依赖于 JDAIS_ASR_IS_NEED_RECORD_FILE 是否使能，如果没有给出录音文件路径，即便开启录音使能，也不会录音
JDAIS_ASR_POST_PROCESS_ENABLE	NSInteger	否	是否开启后处理，0 为不开启，1 为开启，默认为 0 不开启
JDAIS_ASR_PARTIAL_RESULT	NSInteger	否	是否开启实时部分识别结果返回，-1 为不开启，0 为根据服务端配置情况是否开启，1 为开启，默认为 0
JDAIS_ASR_PUNC_PROCESS	NSInteger	否	是否开启识别结果加标点功能，-1 为不开启，0 为根据服务端配置情况是否开启，1 为开启，默认为 0
JDAIS_ASR_VAD_START_DELAY	float	否	VAD 开始延时(单位秒)范围 [0.25,0.5]，默认 0.25
JDAIS_ASR_VAD_END_DELAY	float	否	VAD 结束延时(单位秒)范围 [0.4,2.0]，默认 0.5
JDAIS_ASR_LOCAL_VAD_ENABLE	NSInteger	否	是否启用本地 VAD 截断，0 为不开启，1 为开启，默认为 1 开启
JDAIS_ASR_CONTEXT_ID	NSInteger	否	动态短语 Id
JDAIS_ASR_PHRASE_LIST	JSON String	否	个性话词表短语列表(JSON 数组格式)

语音识别状态回调：

```
@protocol JDAISClientASRDelegate<NSObject>
// event
// 回调事件，请参考 JDAISASRDefines.h 中定义。
// 准备就绪，开始录音：EV_SJDAIS_ASR_SpeechStart
// 用户开始说话：      EV_SJDAIS_ASR_SpeechBegin
// 用户结束说话：      EV_SJDAIS_ASR_SpeechEnd
// 音量回调：          EV_SJDAIS_ASR_SpeechVolume
// 临时识别结果：      EV_SJDAIS_ASR_SpeechParitalResult
// 最终结果：          EV_SJDAIS_ASR_SpeechResult
// result
// 相关回调事件返回携带的结果，JSON 格式类型
// 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
- (void)onEventSpeechRecognition:(NSString *)event
    withResult:(NSString*)result
    withData:(NSData *)data;
```

@end

result 参数 JSON 字符串说明

名称	描述	取值
result		相关回调事件返回携带的结果 JSON 格式类型 EV_SJDAIS_ASR_SpeechStart、 EV_SJDAIS_ASR_SpeechBegin、 EV_SJDAIS_ASR_SpeechEnd 事件没有 result 事件返回。 EV_SJDAIS_ASR_SpeechVolume 对应的 result 格式： { "volume": "20" } EV_SJDAIS_ASR_SpeechParitalResult, EV_SJDAIS_ASR_SpeechResult 格式一致 { "err_code": "0" //识别错误 码, 0: 识别成功, 非 0: 识别失败, 其他错误 信息见错 误码 "request_id": "fe0c804d-dec9-4a80-8167- b730b697ee5b" "err_msg": "speech recognition success" //错误信息 "context": [{ // 识别文本,如果识别失败将 不返回 "text": "背 多 卢比" }] }

Demo 实例

```
// 实例化
JDAISSpeechManager* _JDSpeechManager = [JDAISSpeechManager create:
JDAIS_ENGINETYPE_ASR];
// 注册监听回调
[_JDSpeechManager setListener:self];
// 设置参数
NSString *path =
[[NSSearchPathForDirectoriesInDomains(NSDocumentDirectory,
NSUserDomainMask,
YES) objectAtIndex:0] stringByAppendingPathComponent:@"HandelAudio.pcm"];
_asrSpeechManagerConfig = @{
    JDAIS_ASR_SAMPLE:@(EVoiceRecognizeSampleRate16K),

    JDAIS_ASR_DOMAIN:@(EVoiceRecognizeDomainSearch),

    JDAIS_ASR_SERVER_VAD_ENABLE:@(EVoiceRecognizeVADEnabledNO),
    // APPID 具体值请发送到 jdai-speech@jd.com 进行申
    请
```

```

        JDAIS_ASR_APPID:@"12345678910",
        JDAIS_ASR_RECOGNIZE_TYPE:@(_recognizeType),
        JDAIS_ASR_IS_NEED_RECORD_FILE:@(EVoiceRecogniz
        eRecordAudioNO),
        JDAIS_ASR_RECORD_FILE_PATH:path,

    };

// 启动识别
[self.JDSpeechManager send:JDAIS_ASR_CMD_START
    withParams:[HandelAudioDemoVC
convertToJsonData:_asrSpeechManagerConfig]];

#pragma mark --JDAISClientASRDelegate
- (void) onEventSpeechRecognition : (NSString *)event withResult : (NSString *)result
withData : (NSData *)data {
    __weak typeof(self) weakSelf = self;
    dispatch_async(dispatch_get_main_queue(), ^{
        __strong JDViewController *strongSelf = weakSelf;
        if (strongSelf) {
            NSDictionary *resultDic = [JDViewController
dictionaryWithJsonString:result];
            if ([event isEqualToString: EV_SJDAIS_ASR_SpeechStart]){
                // 准备就绪
            } else if ([event isEqualToString: EV_SJDAIS_ASR_SpeechBegin]) {
                // 开始说话
            } else if ([event isEqualToString: EV_SJDAIS_ASR_SpeechEnd]) {
                // 结束说话
            } else if ([event isEqualToString: EV_SJDAIS_ASR_SpeechVolume]) {
                // 说话音量的返回
            } else if ([event isEqualToString: EV_SJDAIS_ASR_SpeechParitalResult]) {
                // 中间识别结果
            } else if ([event isEqualToString: EV_SJDAIS_ASR_SpeechResult]) {
                // 最终识别结果
            }
        }
    });
}

```

6. 常见问题及错误码描述

错误码	描述	错误原因	备注
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	说话事件超出限制，语音最大长度为60秒	
-1006	没有识别到语音	用户没有说话	
-1007	服务端错误	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 格式语音	
-1010	Appld 鉴权失败	Appld 没有设置或者不合法设置，请发送邮件到 jdai-speech@jd.com 进行申请	
-1011	语音文件头出错	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 要求采样率为 8000/16000 单通道	
-1013	Vad 初始化失败	Vad 资源配置错误	
-1014	京东云网关错误	京东云网关系统错误信息	