

# 捷通华声

# 灵云公有云能力平台

# 4.0.0

# HTTP 开发手册-VPR

灵云科技 源自清华 服务全球

北京捷通华声语音技术有限公司

Beijing InfoQuick SinoVoice Speech Technology Corp.

2015 年 01 月

# 目 录

1. 概述.....	1
2. HTTP 头.....	1
3. VPR 访问接口.....	2
3.1. VPR 注册请求接口.....	2
3.1.1. 访问.....	2
3.1.2. 服务器应答.....	2
3.1.3. 示例.....	3
3.1.4. x-task-config 参数说明.....	3
3.2. VPR 判别请求接口.....	4
3.2.1. 访问.....	4
3.2.2. 服务器应答.....	4
3.2.3. 示例.....	4
3.2.4. x-task-config 参数说明.....	5

## 1. 概述

灵云HTTP-VPR接口，客户端通过HTTP-VPR接口使用POST方式与灵云平台进行通讯，返回数据为XML格式。

## 2. HTTP 头

客户端POST请求时HTTP头信息，必须带上如下：

顺序	名称	值（例子）	含义	备注
1	x-app-key	ac5d5431	应用标识	必选，由捷通公司提供
2	x-sdk-version	3.8	sdk 版本号	必选，使用 http 接口时，可以指定为 3.8
3	x-request-date	2014-6-18 10:10:11	请求时间	必选，每个任务的请求时间不能重复
4	x-task-config	capkey=vpr.cloud.verify , audioformat=alaw	任务参数信息	必选，为 name=value 形式，多个参数以逗号隔开 详情请参考后面的 <a href="#">参数解释</a>
5	x-session-key	18EA9384AA34	请求数据签名	必选 x-session-key 生成算法说明： x-session-key = md5(x-request-date + devkey)
6	x-udid	101:1234567890	设备标识	必选，如使用设备取设备标识号，如不使用设备设置为例子中的默认值
6	x-tid	1234567	天行 ID	必须
7	x-eid		保留字段	保留字段，不用设置

注：头的内容必须为英文。

devkey是应用开发者密钥，在注册开发者用户时由开发者网站统一分配，用作请求数据签名；

## 3. VPR 访问接口

### 3.1. VPR 注册请求接口

本接口在提交音频数据时,同步在 http 请求的回应消息中返回模型注册的结果或模型训练的结果。

#### 3.1.1. 访问

(POST) [ServiceUrl]/vpr/register

说明: 包体中传输音频数据, 包体格式为:

音频段 1 的长度	音频段 1	音频段 2 的长度	音频段 2	... ..
-----------	-------	-----------	-------	--------

音频段长度和音频段必须成对出现。包体中可以包含 1 个或多个(最多 10 个)音频段长度与音频段的组合。

音频段长度字段: 表示紧跟其后音频段的长度。大小为 4 字节, 字节序为 Little-endian。

音频段: 最大为 1MB。

需要以下 HTTP 头: x-app-key、x-sdk-version、x-request-date、x-task-config、x-session-key、x-udid、x-tid。

#### 3.1.2. 服务器应答

应答格式: XML文件

XML的元素(Element)及含义

顺序	名称	值(例子)	含义	备注
1	ResCode	Success/FAILED	完毕/处理失败	必选
2	ResMessage	Success/FAILED	完毕/处理失败	必选
3	ErrorNo	0	返回错误码	必选
4	Result_Token	0E01123456	结果令牌	成功时必选
5	User_Id	123456	用户 ID	第一次注册成功, 必选。

3.1.3. 示例

```
<?xml version="1.0" encoding="utf-8" ?>

<ResponseInfo>

    <ResCode> Success </ResCode>

    <ResMessage> Success </ResMessage>

    <ErrorNo>0</ErrorNo>

    <Result_Token>0E01123456</Result_Token>

    < User_Id>123456</User_Id>

</ResponseInfo>
```

3.1.4. x-task-config 参数说明

以下是有效参数：

✧ capkey（声纹识别能力名称，必选）

Capkey	类型	支持 audioformat
vpr.cloud.verify.telephone	8k16bit 模型，主要应用于电话信道	alaw8k8bit, ulaw8k8bit, pcm8k16bit
vpr.cloud.verify	8k16bit 模型	alaw16k8bit, ulaw16k8bit, pcm16k16bit

✧ user\_id（用户ID）：由数字和字母组成的字符串，长度不超过64个字节。注册时，用户可以不设置，服务会生成user\_id，返回给用户；如果用户设置了，服务会使用用户设置的user\_id，不过，用户必须保证user\_id的唯一性。注册成功后，模型训练，必须设置。

✧ audioformat（音频格式，必选）

有效参数	参数说明
alaw 8k8bit	8k 8bit a-law 数据
ulaw8k8bit	8k 8bit u-law 数据
pcm8k16bit	8k 16bit PCM 音频数据（默认）
alaw16k8bit	16k 8bit a-law 数据

ulaw16k8bit	16k 8bit u-law 数据
pcm16k16bit	16k 16bit PCM 音频数据

## 3.2. VPR 判别请求接口

本接口在提交音频数据和注册模型 ID 时，同步在 http 请求的回应消息中返回音频与模型的匹配结果。

### 3.2.1. 访问

(POST) [ServiceUrl]/vpr/verify

说明：包体中传输音频数据，包体最大 1M。

需要以下 HTTP 头：x-app-key、x-sdk-version、x-request-date、x-task-config、x-session-key、x-udid、x-tid。

### 3.2.2. 服务器应答

应答格式：XML 文件

XML 的元素 (Element) 及含义

顺序	名称	值 (例子)	含义	备注
1	ResCode	Success/FAILED	完毕/处理失败	必选
2	ResMessage	Success/FAILED	完毕/处理失败	必选
3	ErrorNo	4	返回错误码	必选
4	Result_Token	0E01123456	结果令牌	成功时必选
5	Result	YES/NO	音频与模型匹配结果	成功必选

### 3.2.3. 示例

```
<?xml version="1.0" encoding="utf-8" ?>
```

```
<ResponseInfo>
```

```
  <ResCode> Success </ResCode>
```

```
<ResMessage> Success </ResMessage>

<ErrorNo>0</ErrorNo>

<Result_Token>0E01123456</Result_Token>

< Result>YES</Result>

</ResponseInfo>
```

### 3.2.4. x-task-config 参数说明

以下是有效参数：

✧ capkey（声纹识别能力名称，必选）

Capkey	类型	支持 audioformat
vpr.cloud.verify.telephone	8k16bit 模型，主要应用于电话信道	alaw8k8bit, ulaw8k8bit, pcm8k16bit
vpr.cloud.verify	16k16bit 模型	alaw16k8bit, ulaw16k8bit, pcm16k16bit

✧ user\_id（用户ID，必选）：注册成功时，返回的user\_id。

✧ audioformat（音频格式，必选）

有效参数	参数说明
alaw 8k8bit	8k 8bit a-law 数据
ulaw8k8bit	8k 8bit u-law 数据
pcm8k16bit	8k 16bit PCM 音频数据（默认）
alaw16k8bit	16k 8bit a-law 数据
ulaw16k8bit	16k 8bit u-law 数据
pcm16k16bit	16k 16bit PCM 音频数据

✧ threshold（阈值，可选）：服务将音频与指定模型匹配返回相似百分比，记为相似度。相似度 $\geq$ threshold/100时，服务器应答中Result为“YES”； 否则，为“NO”。不设置threshold时，服务使用默认值60。Threshold取值范围：0~100整数。