**长语音实时转写系统接口文档v0.2-带声纹版本**

**长语音实时转写**

**1 接口概述**

实时长语音接口，提供实时文字输出和声纹输出的能力。

**2 接口地址**

ws://10.86.204.80:10800/websocket

**3 接口协议**

**3.1 实时语音识别websocket协议**

**调用流程：**

**3.1.1 通过websocket建立长连接。**

**3.1.2 开始识别时，需要发送text文本消息，这里需要将signal设置为start，同时设置对应的参数。**

**3.1.3 建立连接并且发送start标识，通过websocket发送音频的二进制文本（将音频文件转为二进制文本，可通过ffmpeg转置），rtasr识别后，会将结果返回。**

**3.1.4 结束之后，发送end标识。**

**识别请求**

开始识别：

{

"signal": "start",

"enableItn": true,

"enablePunctuation": true,

"enableVoiceprint": true,

"enableContinuousDecoding": true,

"enablePartialResult": true,

"sampleRate": 16000,

"enableSilenceDetection": true,

"startSilence": 5000,

"endSilence": 1000,

"speechContext": "热词,逗号分隔"

}

结束识别：

{

"signal": "end"

}

请求参数说明：

|  |  |  |  |
| --- | --- | --- | --- |
| 名称 | 类型 | 必填 | 说明 |
| signal | string | Y | start开始识别  end结束识别 |

**识别响应**

{

"type": "partial\_result/final\_result/end\_result",

"start": 0,

"end": 2000,

"name": "说话人",

"content": "识别结果"

}

识别参数说明：

|  |  |  |
| --- | --- | --- |
| 名称 | 类型 | 说明 |
| type | string | partial\_result该类型表示识别的结果不准确，可以丢弃。  final\_result该类型表示识别的结果准确。  end\_result该类型表示识别结束，只有发送了end signal，才会返回该类型结果。 |