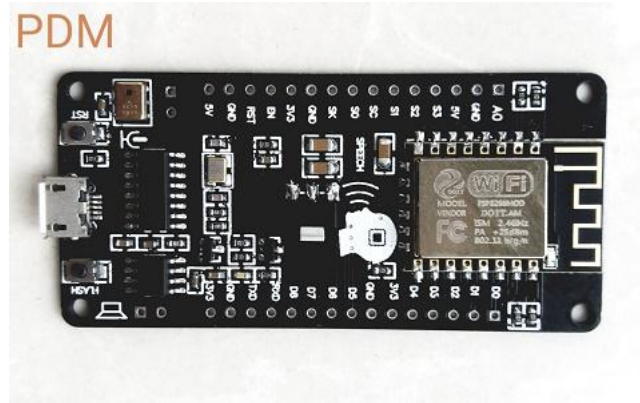


ESP8266 百度云语音识别开发板 SPEECH-ESP 版本说明书



百度云开放了语音识别开发，并且永久免费的，有成熟的 REST API 教学文档，适合爱好者研究，感兴趣者，大家可以到百度学习一下。

我们是基于了乐鑫 ESP8266 的 RTOS SDK2.0 版本的移植。

特点：

1，百度云语音识别开发板子首次加入了 PDM 数字麦克风，仅需要 SCK,DATA2 线，有着极高的信噪比，很高的抗干扰能力，可以紧挨 ESP8266 芯片不受干扰，成功解决了发射时候的“噗噗音”射频干扰声，最大特别的是省去了外置 CODEC 芯片，降低了成本。

2，百度云语音识别开发板，加入了 AT+指令，有着比较详细的功能设置，实现了 AT 指令配置普通话，英语，粤语，四川话，普通话远场的设置，麦克风增益的设置，VAD 设置，WIFI 登陆密码，录音控制，极速版本切换，方便了用户设定语音功能。

3，百度云语音识别开发板，采用了边录制便上传的高效率机制，不用等语音录制结束再上传，这样的目的提高了效率，效果超出了 ESP32 的例子，目前最大上传语音时间是 60 秒，如果上传操作超出这个时间，开发板自动会强制退出上传。识别的所需时间和上传的语音大小有关。

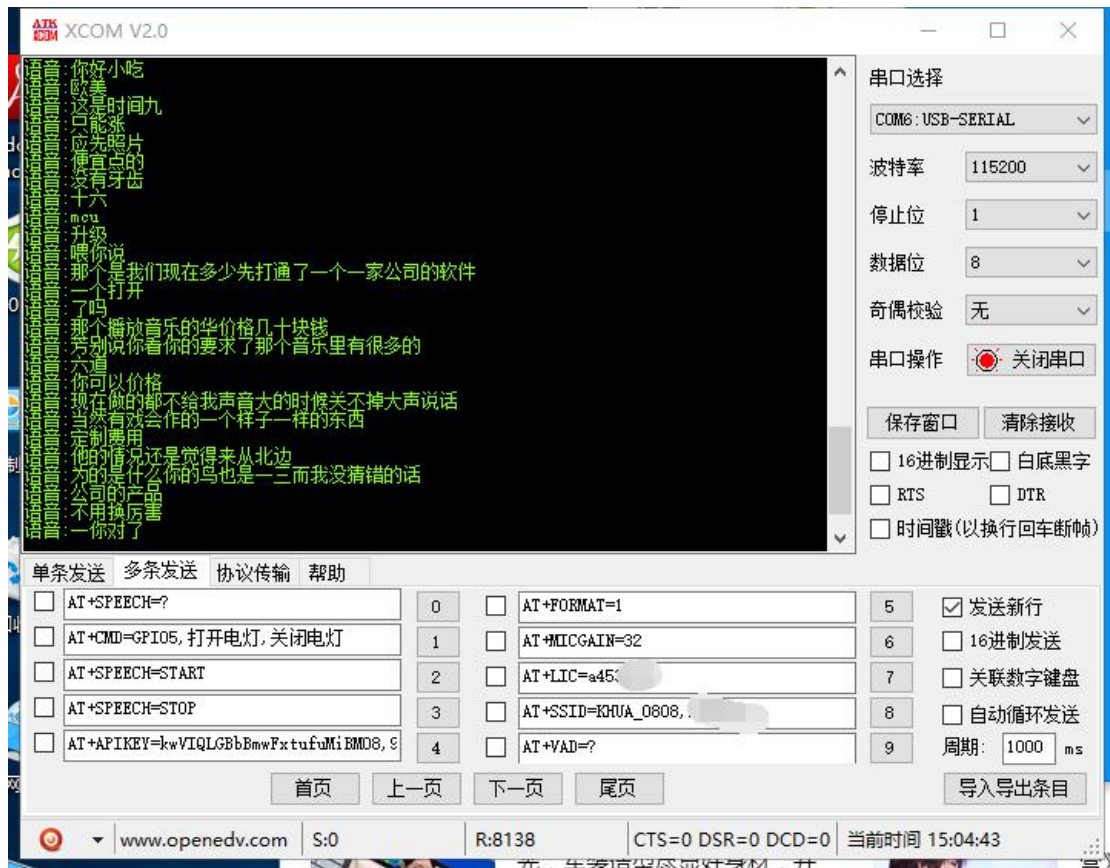
4，百度云语音识别开发板，加入了 VAD 语音检测库，VAD 是一个可以检测语音活动区开始端和语音静音区的结束端的一个工具，可以实现无按键操作，释放了双手。

5，百度云语音开发板，加入了 KWS 唤醒词能力，可以在唤醒词后操作特定命令，更可喜的是，唤醒词可以用户通过 AT 指令设置，并且不用训练，高昂的唤醒词的定制门槛较高，并不适合爱好者学习研究的对象。

6，百度云语音开发板，支持命令词的修改，可以通过 AT 指令设定命令词甚至制定输出引脚的设定。

使用说明：

下载一个串口工具，并且确保安装驱动 CH340G 成功。建议使用 XCOM V2.0。配置 115200,n,8,1.



1, 配置网络, 例如[AT+SSID=SSID,PWD],其中 SSID 是路由器的热点, PWD 是路由器密码。使用方法很容易理解。如果要查询当前 WIFI 配置, 可以使用[AT+SSID=?]可以看到配置的信息了。如果出现 “ip:192.168.1.6,mask:255,255,255,0,gw:192.168.1.1” 类似信息表明配置成功了。

2, 话筒增益配置, 例如[AT+MICGAIN=16],表明配置 PDM 话筒增益为 16, 如果噪音环境较大, 可以适当降低增益, 虽然距离会下降, 但是提高了抗干扰能力, 配置内容范围是 1-32。

3, 唤醒词的设定, 举例[AT+WKS=小智你好],建议这个词一定要容易识别, 熟悉的, 不要用偏僻的字, 最好实际测试一下, 多调整一下, 直到自己满意为止, 举个例子, 如果你设定 “刘德华” 这个唤醒词, 很难识别错误的。注意唤醒词使用必须在 VAD 开启模式下。

4, VAD 的设置, [AT+VAD=ON],配置发送后, VAD 能力开启, 该指令有 ON,OFF, 2 个配置内容, 这时候串口回显 OK,表明开发板支持语音唤醒功能了, 比如说话者对着板子说出 “你好小智” 唤醒词, LED 会闪烁, 表明已被唤醒, 并且在串口会打印 GBK 编码 “语音: 你好小智”, 注意提醒一下, 唤醒后请在是几秒内说出语音命令词, 否则超时后会自动退出唤醒状态。在 VAD 开状态下, 按键操作, AT 操作录音指令是禁止的, 这一点也要注意。

5, 手动录音识别, 执行[AT+VAD=OFF]指令后, 关闭 VAD 功能, 进行[AT+SPEECH=START]开始录音, 并且上传云识别, 这时候 LED 会点亮, 说话者可以不用说唤醒词, 直接讲话说出命令词后, 再次执行[AT+SPEECH=OFF],录音并且上传结束, 并且在还口打印出识别的结果。百度语音识别开发板, 支持手动按键录音识别能力, 空闲模式下, 指的是没有录音等任务的情况下, 按下 FLASH 按键不放手, LED 点亮, 这时候说话者对着开发板直接说话, 说话结束后放手, LED 熄灭, 串口 GBK 编码显示出返回的识别结果。

6, 百度和账号的配置, 百度语音开发板支持百度账号配置, 前提用户要上百度官网注册开发者, 申请到 APIKEY 即可, 申请的详细过程可以百度一下, 申请认证通过后, 可以创建应用, 并且拿到了 API KEY 和 Scret Key 了, 然后按照[AT+=APIKEY=APIKEY,SeretKey]设定发送



并且回传 OK 就可以更新 APIKEY 了。你也可以通过[AT+APIKEY=?]查询账号信息。

7, 百度急速模式配置, 通过指令[AT+SPHMOD=FAST]可以配置极速模式, 响应速度会成倍提高的, 但是遗憾的是有次数限制, 商用的可以考虑, 我们开发者一般设置成 NORMAL 即可。

8, 识别语音模型的配置, 通过指令[AT+DEVPID=1536]配置, 实现了语音的配置。配置的内容有, 1536 普通话模式 (仅仅支持简单的英文识别), 1537 普通话 (纯中文识别), 1637 粤语, 1837 四川话, 1936 远场普通话, 80001(极速版普通话)。

9, 帮助命令[AT+HELP=1], 执行后会看到所有指令及简介。

10, 设备重启[AT+REST=1], 执行后设备会重启。

11, 获取版本信息[AT+VER=?], 可以看到版本信息内容。

12, 语音指令的配置, 例如发送[AT+=GPIO4,打开电灯, 关闭电灯], 其中例子"GPIO4"位置参数目前支持 2 个引脚(GPIO4/GPIO5)内容设定,"打开电灯"参数执行的是引脚输出高电平,"关闭电灯"参数位置识别后执行低电平。举个例子, 如果你将参数改成[AT+CMD=GPIO4,关闭电灯,打开电灯]后, GPIO 引脚正好输出和语义相反的电平。

13, 恢复初始化设备指令[AT+=FORMAT=1]后, 所有信息恢复初始化状态。

14, 技术支持指令[AT_LIC=XXXXXXX], 获取技术支持的注册信息。

未完功能还在陆续增加中.....