

---

# 语音识别芯片使用说明书 V1.02

## 目 录

1、芯片特征.....	2
2、功能描述.....	2
3、应用范围.....	2
4、应用方框图.....	3
5、芯片管脚介绍.....	3
6、芯片电气参数.....	4
7、芯片封装图.....	4
8、芯片功能详解.....	5
8.1 IO 口功能简介.....	5
8.2 指令格式.....	6
8.3 词条相关简介.....	6
8.3.1 词条的存储容量.....	7
8.3.2 词条的存储方式.....	7
8.3.3 词条的识别过程.....	7
8.3.4 词条的设计.....	8
8.3.5 识别成功后的操作.....	8
8.4 识别功能详解.....	9
8.5 播放功能详解.....	9
9、典型应用电路.....	10
10、操作过程.....	11
10.1 串口设置.....	11
10.2 使用方法.....	12
10.3 注意事项.....	13
11、 版本信息.....	13

## 1、芯片特征

- ☆ 半米内识别率 90%以上
- ☆ 可以通过语音指令控制输出
- ☆ 可以通过单片机串口控制播放指定语音
- ☆ 三种输出模式：串口输出、IO 输出、喇叭播放
- ☆ PWM 输出可直接推动 0.5W/8Ω 或 1W/8Ω 扬声器
- ☆ 具有差分放大电路，安静环境下语音识别距离可达到 2-3 米
- ☆ 外挂 SPI Flash，增加存储容量
- ☆ 可以存储多组词条，但是同一时刻只能出现一组词条
- ☆ 非特定人语音识别
- ☆ 语音可支持 32 种语言，目前中文识别度较好

## 2、功能描述

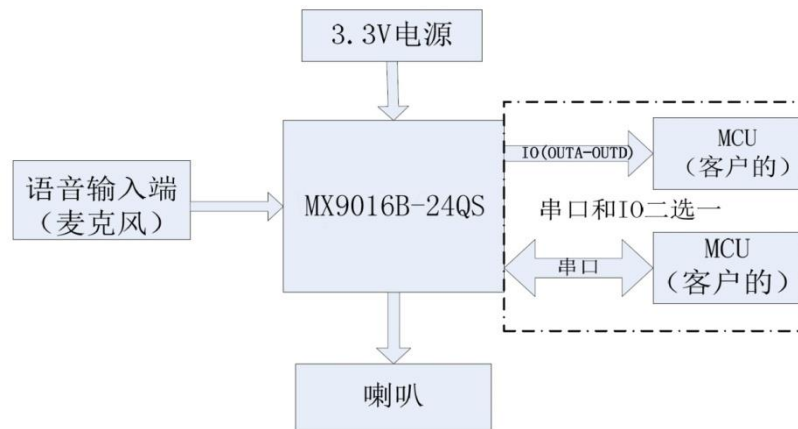
MX9016B 语音识别模块是一款小巧语音识别模块，安静环境下语音识别距离可达到 2-3 米；外挂 SPI Flash, 增加存储容量；主要功能如下：

- ☞ 识别功能：可识别预设的语音词条，同时通过串口，IO 输出以及识别结果。
- ☞ 播放功能：通过单片机指令播放指定的语音，同时通过 BUSY 脚标志是否正在播放语音。

## 3、应用范围

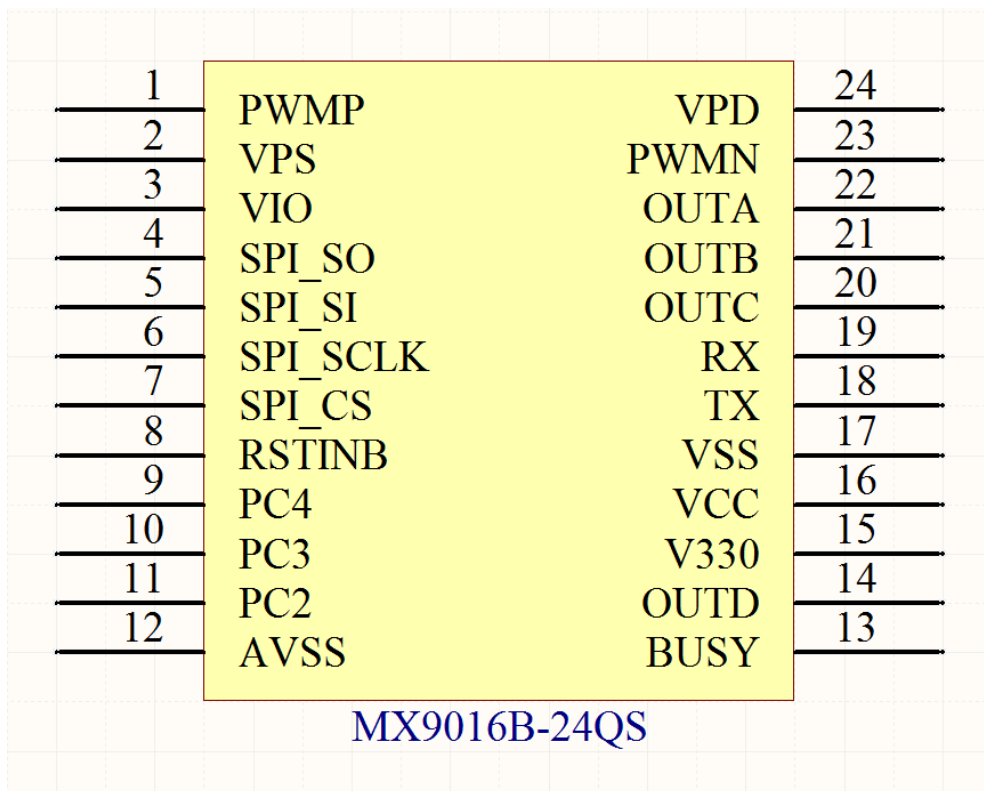
- 玩具类：智能机器人、智能故事机、智能玩具车机等
- 家电类：智能茶壶、智能电饭锅、智能音箱、智能插排、智能晾衣架，智能空气净化器等
- 理疗类：眼部按摩仪、按摩椅、头部按摩仪、足浴按摩器
- 语音导航类：医院、商场等地方内指路器

### 4、应用方框图



结构组成：MX9016B-24QS 语音识别芯片外挂 SPI-flash；可搭配麦克风、喇叭、单片机等外设一起使用。不过，在虚线框中的串口和 IO 都是用于检测 MX9016B 是否能正确识别，所以在连接 MCU（客户的）时可连接其 IO 口或串口，二选一即可。

### 5、芯片管脚介绍



芯片管脚图

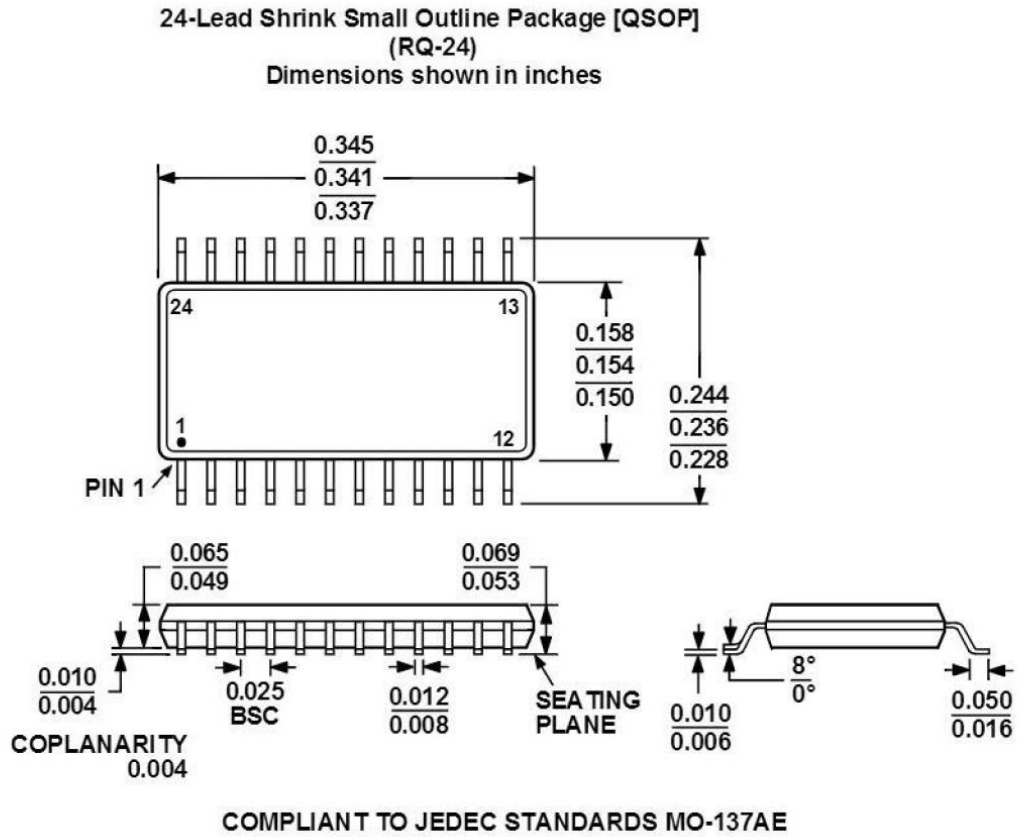
引脚按照上图排列顺序

引脚	引脚标号	功能描述
1	PWMP	数字 PWM 输出(+), 跟 PWMP 组合使用可直接驱动扬声器
2	VPS	PWM 电源输出脚
3	VIO	PortA IO 电源输入脚
4	SPI_S0	SPI_S0 脚
5	SPI_SI	SPI_SI 脚
6	SPI_SCLK	SPI_SI 脚
7	SPI_CS	SPI_CS 脚
8	RESTINB	复位脚 (烧录时需用到)
9	PC4	MIC 麦克风口
10	PC3	MIC 麦克风口
11	PC2	MIC 麦克风口
12	AVSS	模拟地
13	BUSY	BUSY 输出脚
14	OUTD	IO 口/32K 的晶体输出脚
15	V330	内部 LD03.3V/1.8V 电压输出
16	VCC	电源输入脚
17	VSS	数字地
18	TX	串口发送脚
19	RX	串口接收脚
20	OUTC	IO 口/AD 口
21	OUTB	烧录引脚, 闲置时可当 IO 口使用
22	OUTA	烧录引脚, 闲置时可当 IO 口使用
23	PWMN	数字 PWM 输出(-), 跟 PWMP 组合使用可直接驱动扬声器
24	VPD	PWM 电源输入脚

## 6、芯片电气参数

Symbol	Parameter	Min	Max	Type	Unit
VCC	工作电压	1.8	5.5	3.3	V
Iop	无负载电流	--	--	7	mA
Ta	工作温度	-20	75	--	°C
Io	输出电流	-16	16	--	mA

## 7、芯片封装图



## 8、芯片功能详解

### 8.1 IO 口功能简介

- 1) RXD/TXD 标准串口通信、波特率 9600，配置见下图
- 2) 13 脚为播放状态 BUSY 指示，播放时为 H，不播放时为 L
- 3) MIC 麦克风正极输入端，可接常用的麦克风型号

4)PWMP/PWMN pwm 输出、可直接驱动 0.5W/8Ω 或 1W/8Ω 扬声器



串口配置图

## 8.2 指令格式

串口指令格式采用：起始码+参数+反码，设置如下表。

起始码	参数	反码
AA/BB	见下文	见下文

注：①参数+反码 = FF，例如 TX 输出的指令“AA 01 FE”，起始码为 AA，参数为 01，反码为 FE，01+FE=FF，其他数据依此类推

②TX 输出指令起始码为 AA（见表 4），RX 输入指令起始码为 BB（见表 5）

## 8.3 词条相关简介

MX9016B 语音识别模块可以实现固定词汇，非特定人的语音识别。

固定词汇并非指词条永远无法改变，而是用户不可以自行更改，但开发者可以根据需求更换词条，之后将词条存储到 MX9016B-24QS 语音芯片里或是外挂的 SPI Flash 里，待识别时将会拿输入的语音与存储起来的词条语音作对比，若是两者相似度达到标准就认为是输入的语音与存储中的词条一致，即识别成功。

### 8.3.1 词条的存储容量

存储词条的地方有两处：语音芯片的内部、SPI Flash;两者本身的容量不同，所能存储的词条数量自然也不同。下面进行四字词条（即是每个词条四个字）的存储容量的对比

存储位置	存储词条数量
MX9016B-24S 语音芯片	12 条
SPI Flash	根据 Flash 自身容量决定

虽说 MX9016B-24QS 语音芯片内部可以存储 12 条左右的词条，但是存储的数量越多，运算的速率越低，识别的效果也会相对的差一点。

### 8.3.2 词条的存储方式

词条的存储方式主要有两种：

#### ①一组存储

在词条的数量较少，并且都是同一类型，如都是表示颜色的词条，那可以选择放在同一组中，不用进行切换，更省时间。

#### ②分组存储

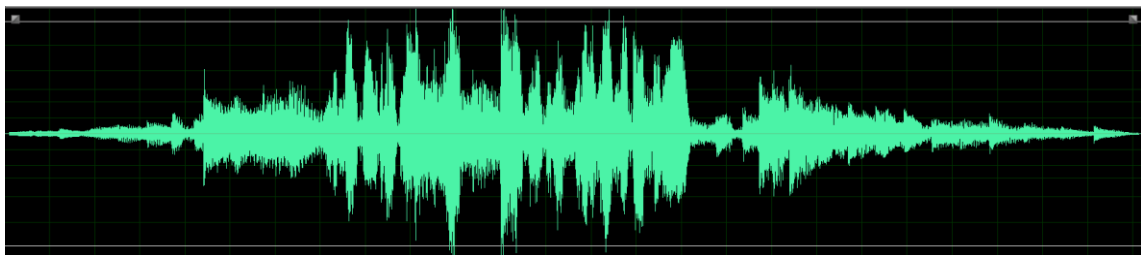
若是词条的数量多，并且分类不一样，那可以选择分组存储，例如现在有三组不同类型的词条，分别是故事类，古诗类以及音乐类

故事类	古诗类	音乐类
拔苗助长	望天门山	爱我中华
后羿射日	游山西村	光辉岁月
亡羊补牢	使至塞上	好久不见
守株待兔	过零丁洋	心如刀割

将词条分组存储，待需要进行古诗类词条时，则需先进入古诗类词组，那么我们需要一个入口。

### 8.3.3 词条的识别过程

词条的识别过程实际就是波形的识别。实际上，声音会产生波形，可见如下声音波形图。无论是人声，还是风扇的声音，都有波形，但是不同的人说同一句话，声音的波形都会有一点点的差别，就算是同一个人说同一句话，声音的波形也会有差别，因为音调，语速等等都可能发生细小的变化，从而影响声音波形。



因此，在设计的过程中，会设置一个门槛，若输入语音的波形与存储好的词条波形相似度高于这个门槛，则识别成功。而这个门槛的高低可以改动的（用户无法更改），门槛调高了，即相似度要更高，会更难识别，识别率降低；若门槛调低了，更容易识别，识别率增加，但是这也带来了一个问题----误识别。

在吵杂的环境中，多种声音的波形叠加在一起，刚好与存储中的某一个词条的波形相似度超过门槛值，那么就会认为是存储中的词条，从而语音识别成功，这就引起了误识别，若设置的门槛越低，误识别率就越高。

所以说，识别率和误识别率是一致的，识别率升高，误识别率也会升高，这就需要根据场合，根据需求来设计识别率。当然，减低误识别率是有办法的，如在不需要识别的时候，让模块处于休眠状态，既可以省电，又可以降低误识别率。

### 8.3.4 词条的设计

识别率，除了与识别的门槛有关，还与词条的长短等有关，其关系见表 6。

词条与识别效果表

词条	识别效果
开	太短了，不好识别
开灯	识别效果一般
关灯	与其他词条相似，识别效果一般
小智，开灯	长度适中、相似度低，识别效果佳
我困了，请关灯	长度适中、相似度低，识别效果佳

因此，在更换词条时，应考虑词条的长度，各个词条间的相识度，尽量做到长度适中、相似度低。

### 8.3.5 识别成功后的操作

MX9016B 语音识别模块具有三种输出模式：串口输出、IO 输出、喇叭播放；我们可以根据需要选择一种方式，或者两种、三种方式，检测语音是否识别成功。而喇叭播放的内容也是可以根据需要更改的（用户不可自行更改）。



## 8.4 识别功能详解

从麦克风中输入词条（即语音指令），与预置词条的特征因子进行匹配，如果一致，即识别正确，之后模块的串口 TX 会输出相应的指令，并且三个 IO 口（OUTA-OUTC）会输出相应的电平，同时喇叭播放语音。以下是一个工程案例，词条及应答语需要根据每个案子定制，固化在芯片。

词条、TX 串口指令及 IO 电平表

IO 端子命令	OUTC	OUTB	OUTA	TX 输出	执行结果
XX 你好	0	0	0	AA 00 FF	语音播放主人您好
请开灯	0	0	1	AA 01 FE	第一个灯亮，语音播放灯已打开
白色光	0	1	0	AA 02 FD	第二个灯亮，语音播放灯已调至白色光
自然光	0	1	1	AA 03 FC	第一、二个灯亮，语音播放灯已调至自然光
黄色光	1	0	0	AA 04 FB	第三个灯亮，语音播放灯已调至黄色光
银河夜灯	1	0	1	AA 05 FA	第一、三个灯亮，语音播放灯已调至银河模式
助眠夜灯	1	1	0	AA 06 F9	第二、三个灯亮，语音播放灯已调至助眠模式
最大亮度	1	1	1	AA 07 F8	三个灯全部亮，语音播放灯已调至最大亮度
一半亮度	0	0	0	AA 08 F7	三个灯都不亮，语音播放灯已调至一半亮度

注：“1”表示高，“0”表示低

## 8.5 播放功能详解

播放功能是与 MCU（用户的）共同实现的，通过 MCU 来播放指定的语音（单独使用语音识别模块也可实现此功能）。

上电后，MCU（用户的）通过串口给 MX9016B 语音识别模块发送指令，播放 SPI Flash 中的音频，如果成功播放，MX9016B 语音识别模块会通过串口输出“01”，如果因输入指令有误，无法播放则输出“02”，并且在播放的时候会将 BUSY 脚的电平拉高，等播放完毕或是无法播放时将 BUSY 脚的电平拉低。因此，根据 BUSY 脚可计算语音播放的时间。具体的指令及 BUSY 脚电平情况见下表。

播放功能所需数据表

RX 输入	播放SPI Flash 的音频 (SPI 音频可修改)	成功播放 (TX 输出返回指令)	输入有误无法播放 (TX 输出返回指令)	BUSY (播放时)	BUSY (播放结束)
BB 01 FE	SPI Flash 第一段语音	01	02	1	0
BB 02 FD	SPI Flash 第二段语音	01	02	1	0
BB 03 FC	SPI Flash 第三段语音	01	02	1	0

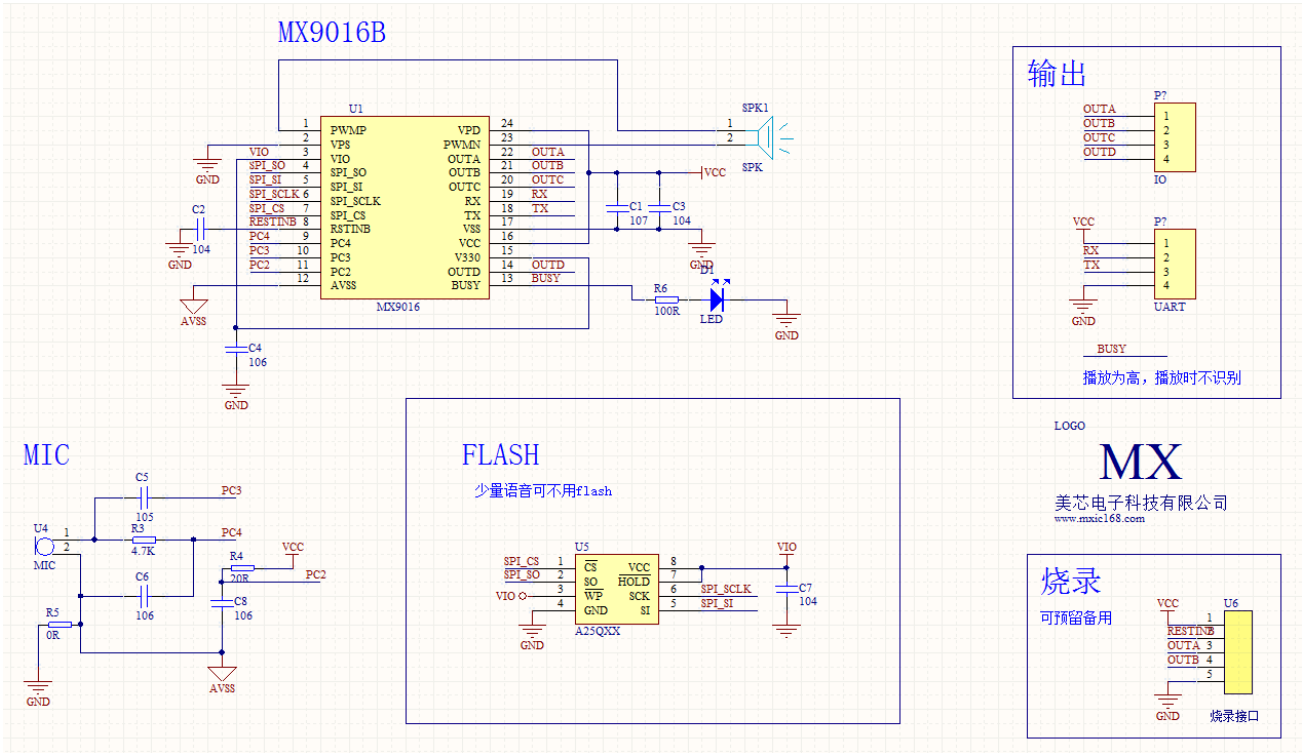
BB 04 FB	SPI Flash 第四段语音	01	02	1	0
BB 05 FA	SPI Flash 第五段语音	01	02	1	0
BB 06 F9	SPI Flash 第六段语音	01	02	1	0
BB 07 F8	SPI Flash 第七段语音	01	02	1	0
BB 08 F7	SPI Flash 第八段语音	01	02	1	0
BB 09 F6	SPI Flash 第九段语音	01	02	1	0
BB 0A F5	SPI Flash 第十段语音	01	02	1	0
BB 0B F4	SPI Flash 第十一段语音	01	02	1	0
BB 0C F3	SPI Flash 第十二段语音	01	02	1	0

注：“1”表示高，“0”表示低

## 9、典型应用电路

电路中接入麦克风当语音的输入端，而喇叭、串口、IO 作为输出端。

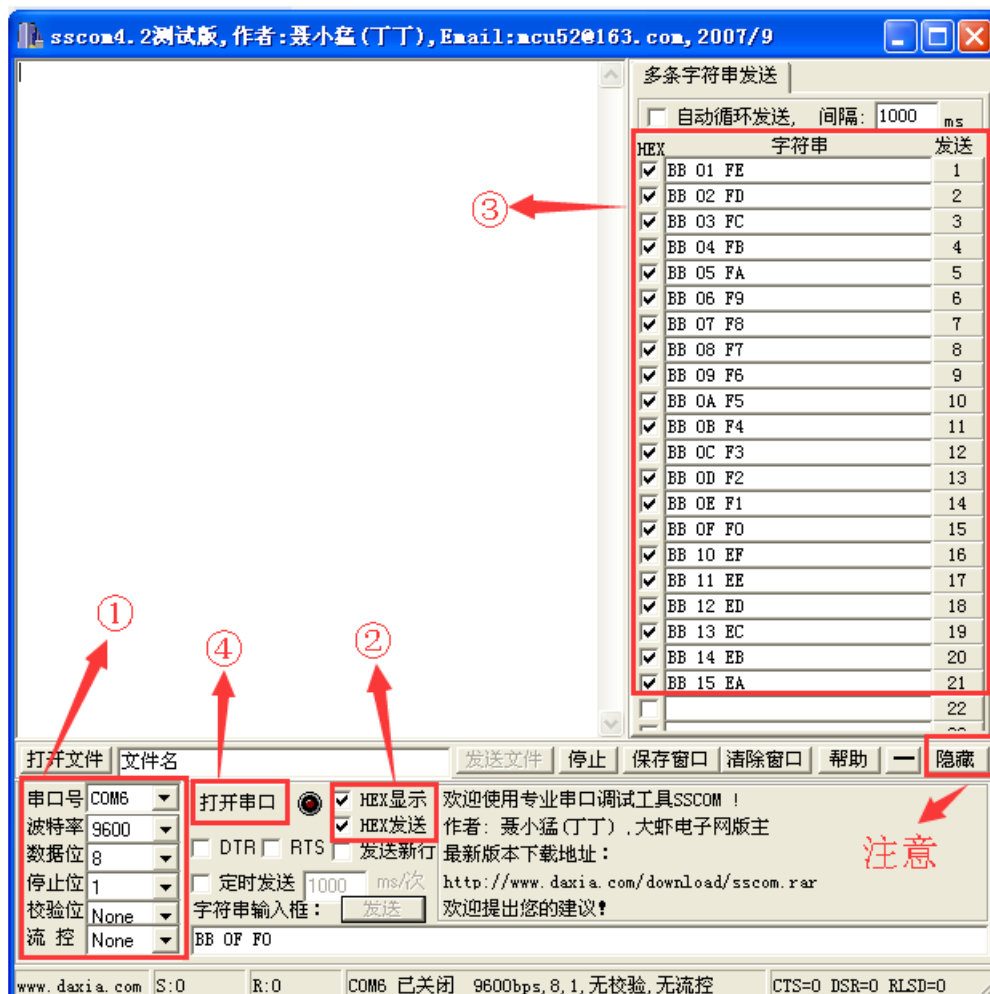
- 1) 需要播放语音时则需接入喇叭
- 2) 播放功能中，MX9016B 语音识别模块与客户的 MCU 相连，可连接 MCU 的 IO 口，也可连接 MCU 的串口，二选一即可。



## 10、操作过程

### 10.1 串口设置

在没有连接单片机时，可通过串口调试助手模拟单片机的操作，则需对串口进行相应的设置，具体见下图。



注：如果用户打开软件不出现③的输入框，请点击“注意”中“隐藏”所在位置的按钮，此时按钮为“扩展”，点击后，按钮变为“隐藏”。

#### ➤ 设置①中的参数如下：

串口号为串口板当前使用的端口（用户需根据实际使用端口设置）波特率：出厂默认值为 9600，如客户曾有修改，以修改的值数据位：8 停止位：1 校验位：None 流控制：None

- 设置完毕后，勾选②，即 HEX 显示，HEX 发送，并请检查当前串口端口不被其他设备占用，点击④“打开串口”
- ③中的指令对应播放功能中的 RX 输入指令，可直接将 8.4 节中的表 5 中的 RX 输入列下的指令复制进去，并将其指令前的 HEX 选项勾选上，点击指令后的数字即可发送此指令

## 10.2 使用方法

根据接线步骤、串口设置准备好后，即可使用（上电进入工作模式）。

- ◆ 若是想进行语音识别，只需对着麦克风念出词条，如“请开灯”，待语音识别正确，IC 会播放语音，I/O 口电平发生变化、同时串口输出相应的指令，更多词条识别后的结果见 8.3 识别功能详解中表 4。
- ◆ 若是想通过单片机控制模块，令其播放相应的语音，则通过串口发送指令，按照指令，模块会播放相应的。

## 10.3 注意事项

环境：安静的环境

距离：半米效果最佳，距离越远，识别率会有所下降；因距离越远，声音越难传播到麦克风中心词条时正对着麦克风，可让麦克风更好的接收声音，识别效果更好

念词条时，只需按平常坐姿说话，不需特意将头靠近麦克风

说话的语速适中，忌太快，太慢，用平常说话的速度即可

最好将模块悬空，若是麦克风的焊接处用手压着或是压在可以导电的物体上，会令其认为有波形输入，从而影响识别率

若接上喇叭播放语音，如果想语音更大声，更清晰需接上腔体，防止播放的语音散开，可将喇叭放在一个小盒子里，或是用手圈着喇叭。

## 11、 版本信息

时间	版本号	升级记录
2017-12-12	V1.00	原始版本
2018-01-01	V1.01	修改识别词条、更换管脚图
2018-04-03	V1.02	更改无负载电流