

VTX316中文语音合成芯片_用户手册

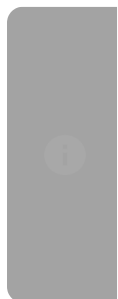
功能特别注意



注意：

本型号芯片：单次支持4000字节的文本合成

外链说明文档



[《文字编码说明》](#) 

[《串口命令说明》](#) 

[《文本控制标记说明-中文合成》](#) 

[《功放芯片推荐参考手册》](#) 

1. 概述

本语音合成芯片是北京宇音天下科技有限公司推出的一款性/价比更高，效果更自然的一款高端语音合成芯片。本芯片通过UART接口，接收待合成的文本数据，实现文本到语音（或TTS语音）的转换。

本语音合成芯片，小尺寸贴片封装、硬件接口简单、低功耗、音色清亮圆润、极高的性/价比；除此之外，芯片在识别文本/数字/字符串更智能、更准确，语音合成自然度上升了一个大的台阶。语音合成效果和智能化程度达到了质的飞跃，是一款真正面向高端行业应用领域的语音合成芯片。

本语音合成芯片的诞生，将推动TTS语音合成技术的行业应用走向更深入、更广泛！



2. 主要应用领域

- 车载终端 智慧停车
- 考勤终端 公交报站

- 排队叫号 预警终端
- 仪器仪表 智能玩具
- 语音对讲 自助终端

3. 产品功能描述

文本合成功能

清晰、自然、准确的中文语音合成效果。芯片支持任意中文文本的合成，可以采用GB2312、GBK、Unicode和UTF8四种编码方式。

单次支持4000字节的文本合成。

芯片支持英文字母的合成，遇到英文单词时按字母方式发音。



详见 [《文字编码说明》文档](#) 

文本智能分析处理

芯片具有文本智能分析处理功能，对常见的数值、电话号码、时间日期、度量衡符号等格式的文本，芯片能够根据内置的文本匹配规则进行正确的识别和处理。

- 例如：“2012-05-01 10:36:28”读作“二零一二年五月一日十点三十六分二十八秒”，“火车的速度是622km/h”读作“火车的速度是六百二十二公里每小时”，“-12°C”读作“零下十二摄氏度”，等等。

多音字和姓氏处理

对存在多音字的文本，例如：“银行行长穿过人行道向骑着自行车的银行职员行走过去”，芯片可以自动对文本进行分析，判别文本中多音字的读法并合成正确的读音。

有些汉字作为姓氏使用时，会出现一些非常规的读法，芯片可以自动进行处理。例如：“他是一位姓朴的朴素的韩国艺人。”，句中两个“朴”字前面一个读作“piao2”，后面一个读作“pu3”。

音量、语速、语调调整

芯片可实现10级数字音量控制，音量更大，更广。支持30级语速、10级语调的调节，满足各种不同的应用需求。

提示音

芯片内集成了多首声音提示音，可用于不同行业不同场合的信息提醒、报警等功能。



详见 [《提示音》](#) 章节

支持多个发音人

提供两男、两女、一个效果器和一个女童、两个男童声共8个中文发音人，可以通过使用特殊标记[m?]来切换芯片的发音人。

发音人标记	说明
[m3]	女声“晓玲”
[m51]	男声“尹小坚”
[m52]	男声“易小强”
[m53]	女声“田蓓蓓”
[m54]	效果器“唐老鸭”
[m55]	女童声“小燕子”
[m56]	男童声“贝童”
[m57]	男童声“晓可”

多种文本控制标记

芯片支持多种文本控制标记。可通过发送“合成命令”发送文本控制标记，调节语速、语调、音量。

还可以使用控制标记提升文本处理的正确率，如：设置句子的韵律、设置数字读法、设置姓氏读音策略、设置号码中“1”的读法等。



详见 [《文本控制标记说明-中文合成》文档](#) 

多种控制命令

控制命令包括：合成文本、停止合成、暂停合成、恢复合成、状态查询、进入Deep Sleep模式。控制器通过通讯接口发送控制命令实现对芯片的控制。

查询芯片的工作状态

支持多种方式查询芯片的工作状态，包括：查询状态管脚电平、通过读芯片自动返回的回传、发送查询命令获得芯片工作状态的回传。

通讯模式

通讯模式：芯片支持UART通讯方式。

支持低功耗模式

芯片支持 Deep Sleep模式。使用控制命令可以使芯片进入 Deep Sleep模式。

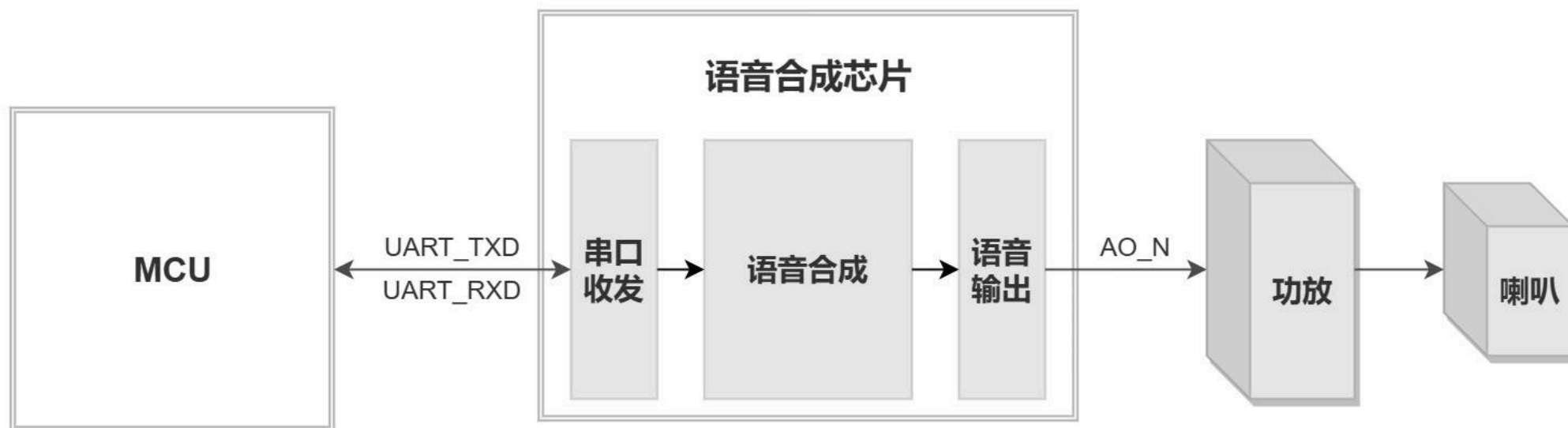
支持 4 种通讯波特率

芯片支持的通讯波特率：9600bps、57600bps、115200bps 、460800 bps。

4. 系统构成框图

最小系统包括：控制器模块、语音合成芯片、功放模块和喇叭。

主控制器和语音合成芯片之间通过UART 接口连接，控制器可通过通讯接口向语音合成芯片发送控制命令和文本，语音合成芯片把接收到的文本合成为语音信号输出，输出的信号经功率放大器进行放大后连接到喇叭进行播放。

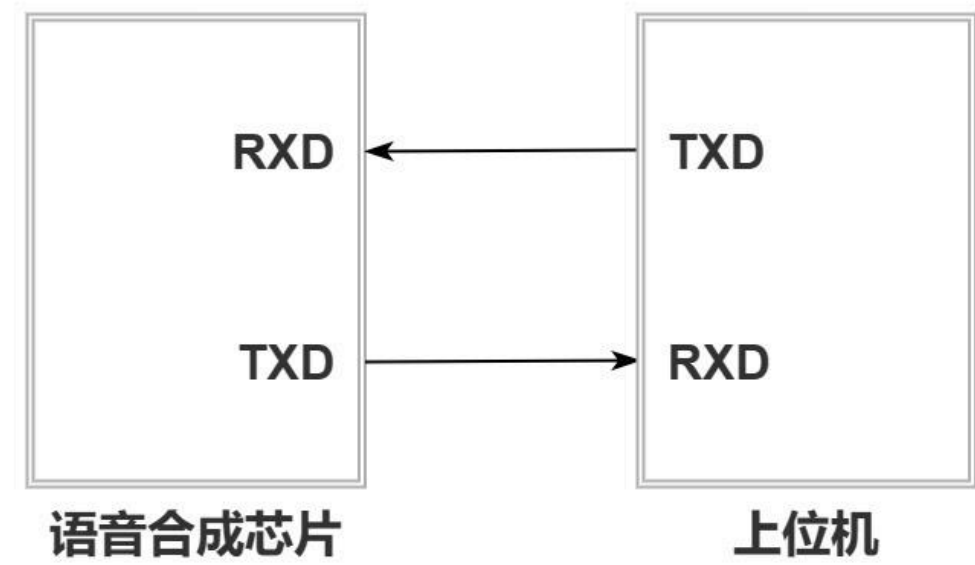


5. 芯片通讯方式

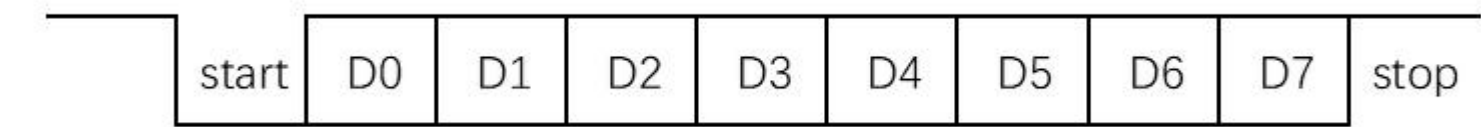
芯片支持UART接口通讯方式，可通过UART 接口接收上位机发送的命令和数据。

硬件连接

本芯片提供一组全双工的异步串行通讯（UART）接口，实现与微处理器或PC的数据传输。芯片利用TxD和RxD以及GND实现串口通信。其中GND作为信号地。芯片支持UART 接口通讯方式，通过UART 接口接收上位机发送的命令和数据。



通讯传输字节格式



通讯标准	波特率	起始位	数据位	停止位	校验
UART	9600bps、57600bps、115200bps、460800bps	1bit	8bits	1bit	无

波特率配置方法

本芯片的UART 通讯接口支持4 种通讯波特率：9600bps，57600bps，115200bps，460800bps，可以通过芯片上的两个管脚BAUD0、BAUD1上的电平来进行硬件配置，两个IO默认上拉，接10K电阻，如果悬空为115200bps。

0：表示低电平 1：表示高电平

波特率	BAUD0	BAUD1
9600 bps	0	1
57600 bps	1	0
115200 bps	1	1
460800 bps	0	0

6. 串口命令说明

上位机发送给芯片的所有命令和数据都需要用“帧”的方式进行封装后传输。

详见 [《串口命令说明》文档](#) 

注意：

本芯片的“特殊参数配置”默认配置为：

- 1、“去上电POP音延时”=0，最大可调整到200，200=200ms；
- 2、“去句首丢音延时”=0，最大可调整到250，250=250ms；
- 3、“去句尾POP音延时”=50，最大可调整到300，300=300ms。

7. 控制标记&合成参数

为满足客户在各种复杂环境的个性化需求，特推出以下功能：

文本控制标记： 用于语音合成参数的设置，并可以灵活控制汉字发音的变化（例如：设置音量，设置语速，设置发音人，强制某个汉字的发音，强制某串数字按号码还是按数值发音等）

语音合成参数设置： 发送相应的“文本控制标记”即可进行语音合成参数的设置



详见 [《文本控制标记说明-中文合成》文档](#)

8. 提示音

提示音				
序号	名称	播放时间	声音类型	试听
1	sound101	0.8秒	警报	0:00 / 0:00
2	sound102	0.16秒	刷卡成功	0:00 / 0:00
3	sound103	2.4秒	风铃声	0:00 / 0:02
4	sound104	0.09秒	刷卡成功	0:00 / 0:00

5	sound105	0.11秒	刷卡成功	0:00 / 0:00
6	sound106	1.1秒	警报	0:00 / 0:01
7	sound107	0.8秒	警报紧急	0:00 / 0:00
8	sound108	1.2秒	叮咚	0:00 / 0:01
9	sound201	1.3秒	叮咚	0:00 / 0:01
10	sound202	1.9秒	警报声	0:00 / 0:01
11	sound901	0.05秒	嘀（蜂鸣器）	0:00 / 0:00
12	sound902	0.14秒	嘀（蜂鸣器）	0:00 / 0:00

13	sound903	0.87秒	咚	0:00 / 0:00
----	----------	-------	---	-------------

举例：

`sound901[p100]sound901[p100]sound901[p100]sound901`

说明：蜂鸣4次，每次蜂鸣持续50ms，两次蜂鸣间隔100ms，*[p100]=间隔100ms*

8.1. 提示音注意事项

提示音在使用上没有特殊性，与合成普通文本的合成命令相同。

但是，需要注意的是：提示音名称前面或后面紧接着是英文字母和数字串时，需要使用标点符号、空格、回车等与其他字母隔开，系统才能够自动识别。

例如：发送文本“sound901, hello! ”，sound901 就可以合成对应的蜂鸣器提示音，但是如果发送的文本“sound901hello! ”，sound901 就不能合成提示音，而是直接朗读成字母“S-O-U-N-D”。

如果不是上述的有效提示音，则只会按着普通文本发音

连续播音实现

我们可以通过多次播放同一提示音来实现一个较长提示音的效果

如需提示音之间有停顿感，可在提示音中间增加加 “[p?]” 或 “逗号” 来实现

举例如下：

1. sound901[p100]sound901[p100]sound901[p100]sound901
2. sound901sound901sound901sound901sound901, sound901sound901sound901sound901sound901, sound901sound901sound901sound901sound901
3. sound901, sound901, sound901, sound901, sound901, sound901, sound901, sound901, sound901, sound901, sound901, sound901, sound901, sound901, sound901
4. [p20]sound901sound901sound901sound901, [p20]sound901sound901sound901sound901, [p20]sound901sound901sound901sound901, [p20]sound901sound901sound901sound901

举例1：蜂鸣4次，每次蜂鸣持续50ms，两次蜂鸣间隔100ms

举例2：每播放5个提示音会有一个停顿

举例3：每播放1个提示音会有一个停顿

举例4：每播放4个提示音会有一个20ms的停顿

以上几种方式都可实现提示音连播，用户可根据实际需求自行选择，但需注意：



1、连续字符（含字母和空格等）每当超过60个字符前必须加个逗号，否则提示音会被中途切割不能连续播放

2、每个提示音的实际播放时间=主体音“播放时间”+0.1秒尾部静音。例如：“sound901”的“播放时间”是0.05秒，实际播放时间是0.15秒；“sound901sound901”的实际播放时间是0.30秒。

9. 合成示例程序

9.1 C语言范例

下面以 51 单片机为作为上位机为例，用C51 语言实现一段文本合成的程序实例，假设要合成的文本内容为：“欢迎使用宇音天下研发的语音合成芯片”，以下为发送一帧TTS 文本数据的程序模块。

```
1  #include <reg51.h>
2  #include <string.h>
3
4  void main(void)
5  {
6  /*****需要发送的文本*****/
7      char code text[ ] = {"欢迎使用宇音天下研发的语音合成芯片"};
8      unsigned char headOfFrame[5];
9      unsigned char length ;
10     unsigned int i = 0;
11     length = strlen(text);          //需要发送文本的长度
12
13 /*****串口的初始化*****/
14     TL1 = 0xFA;          // 在11.0592MHz 下，设置波特率9600bps，工作方式2
15     TH1 = 0xFA;
16     TMOD = 0x20;
17     SCON = 0x50;          // 串口工作方式1，允许接收
18     PCON = 0x80;
19     EA = 0;
20     REN = 1;
21     TI = 0;              //发送中断标志位置零
22     RI = 0;              //接收中断标志位置零
23     TR1 = 1;             //定时器1 用作波特率发生
```



```

24
25 /*****发送过程*****/
26     headOfFrame[0] = 0xFD      //构造帧头FD
27     headOfFrame 1  = 0x00      //构造数据区长度的高字节
28     headOfFrame 2  = length + 2 //构造数据区长度的低字节
29     headOfFrame 3  = 0x01      //构造命令字：合成播放命令
30     headOfFrame 4  = 0x01      //构造命令参数：编码格式为GBK
31
32     for i = 0  i<5  i++        //依次发送构造好的5个帧头字节
33
34         SBUF = headOfFrame i
35         while  TI== 0          //等待发送中断标志位置位
36             TI = 0             //发送中断标志位清零
37
38
39     for i = 0  i<length  i++    //依次发送待合成的文本数据
40
41         SBUF = text i
42         while  TI== 0
43             TI = 0
44
45
46     // while(1);
47

```

9.2 汇编语言范例

下面是 51 单片机为控制上位机的汇编语言示例，演示向芯片发送“宇音天下”这段文本进行合成。

“宇音天下”的GBK码是：

- “宇”：0xd3ee
- “音”：0xd2f4
- “天”：0xccec
- “下”：0xcfc2

```

1      ;晶振11.0592MHz
2
3      ORG 0030H
4  TABLE: DB 0xD3,0xEE,0xD2,0xF4,0xCC,0xEC,0xCF,0xC2      ;"宇音天下"的GBK码
5
6      ORG 0000H
7      LJMP MAIN
8
9
10     ORG 0200H
11  MAIN:
12     CLR EA      ;串口初始化
13
14     MOV TMOD, #20H      ;定时器1工作在方式2
15     MOV TH1, #0FAH      ;装载定时器初值，波特率9600
16     MOV TL1, #00H
17     SETB TR1      ;启动定时器1
18
19     MOV SCON, #50H      ;串口工作方式1，允许接收
20     MOV PCON, #80H      ;波特率倍增
21     CLR TI      ;发送中断标志位置零
22     CLR RI      ;接收中断标志位置零
23
24     MOV A, #0FDh      ;帧头FD
25     MOV SBUF, A      ;串口发送
26     JNB TI, $

```

```

27      CLR TI
28
29      MOV A, #00h      ;数据区长度的高字节
30      MOV SBUF, A
31      JNB TI, $
32      CLR TI
33      MOV A, #0Ah      ;数据区长度的低字节
34      MOV SBUF, A
35      JNB TI, $
36      CLR TI
37
38      MOV A #01h      命令字：合成播放命令
39      MOV SBUF A
40      JNB TI $
41      CLR TI
42
43      MOV A #01h      命令参数：编码格式为GBK
44      MOV SBUF A
45      JNB TI $
46      CLR TI
47
48      MOV R7 #8        待播放文本的总字节数
49      MOV R6 #0        发送字节计数
50      MOV DPTR #TABLE
51  LOOP: MOV A R6
52      MOVC A @A+DPTR
53      MOV SBUF A      播放"字音天下"
54      JNB TI $
55      CLR TI
56      INC R6
57      DJNZ R7 LOOP
58
59      SJMP $
60  END

```

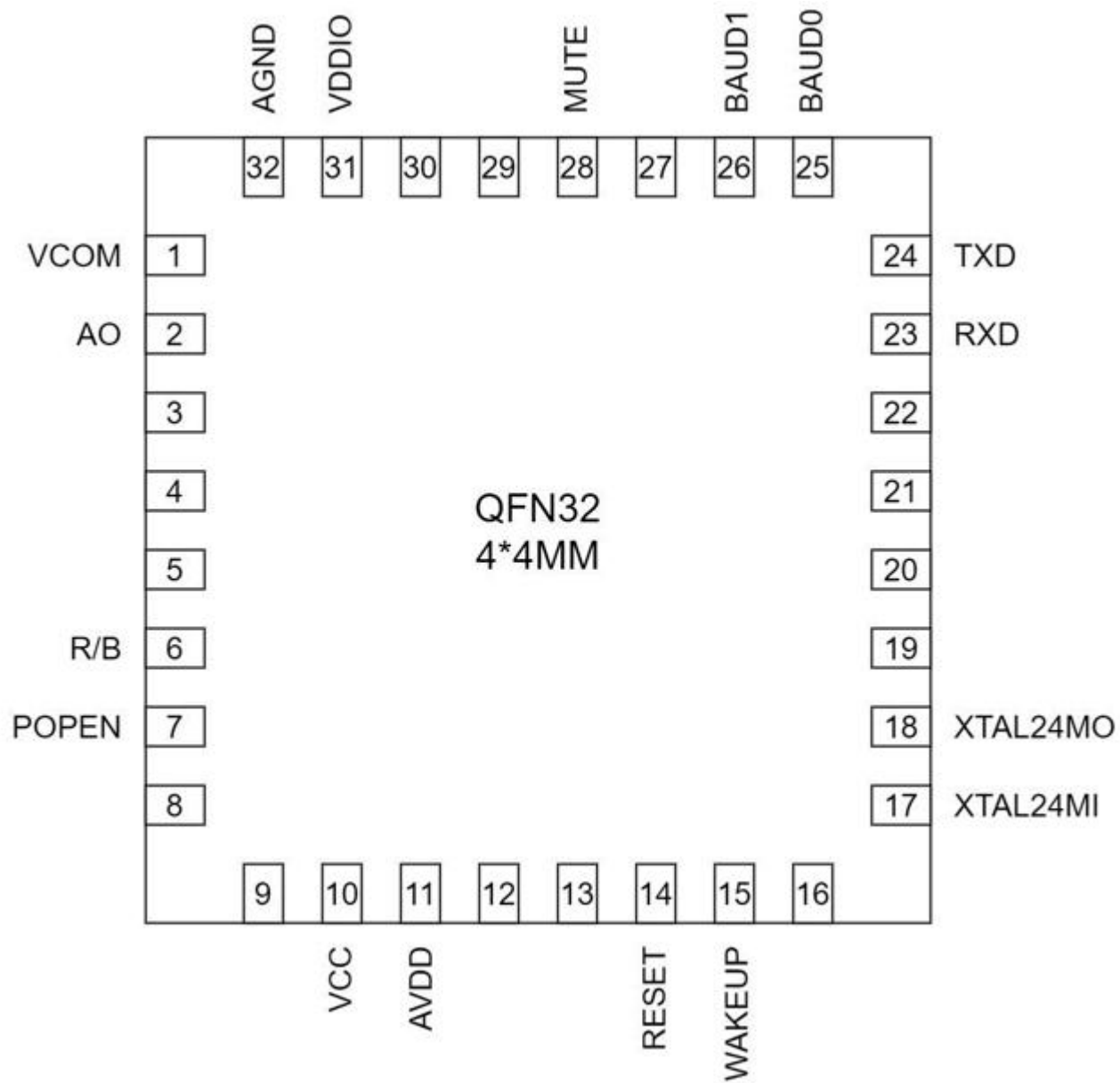
注意：

- 1、发送后可接收到芯片的反馈信号，收到“41”，表明文本被正确接收；收到“4F”，说明合成播音完毕，芯片处于空闲状态；如果收到的是“45”，则说明文本没有正确的收到或是合成，则需要重新发送或者复位。
- 2、上面的Demo 主要说明了发送过程需要遵循的协议问题；发送完语句后要加入发送是否完成的判断程序，可以通过查询或者中断两种方式进行判断当前芯片的工作状态，然后才能发送下一个数据。

10. 芯片引脚

10.1 芯片引脚图

封装名称	封装描述
QFN32	32脚，芯片尺寸4mm×4mm×0.80mm



10.2 芯片引脚定义

编号	引脚	说明	编号	引脚	说明
1	VCOM	接一个1uF电容到AGND	17	XTAL24MI	24M晶振输入
2	AO	音频输出	18	XTAL24MO	24M晶振输出
3	NC	-	19	NC	-
4	NC	-	20	NC	-
5	NC	-	21	NC	-
6	R/B	芯片忙状态指示，低电平表示空闲	22	NC	-
7	POPEN	默认为高电平，详见 POPEN+MUTE口方案	23	RXD	串口接收
8	NC	-	24	TXD	串口发送
9	NC	-	25	BAUD0	波特率配置端口0
10	VCC	芯片供电 2.6-5V	26	BAUD1	波特率配置端口1
11	AVDD	接一个1uF电容到GND	27	NC	-

12	NC	-	28	MUTE	连接功放的静音端口，功放类型是高电平工作 低电平停止工作，详见 POPEN+MUTE口方案
13	NC	-	29	NC	-
14	RESET	复位脚（保持1秒以上低电平后拉高复位）	30	NC	-
15	WAKEUP	睡眠时下降沿唤醒	31	VDDIO	接一个10uF电容到GND
16	NC	-	32	AGND	模拟地

11. POPEN+MUTE口方案

11.1 方案选择

用户根据自身情况合理选择方案，正确配置合成芯片的POPEN口（第7脚）与MUTE口（第28脚）。

方案	POPEN口与MUTE口配置	芯片功耗
方案一	POPEN口：低电平(接1K电阻到GND) MUTE口：接功放使能口	睡眠：不支持低功耗 启动：5mA 空闲：7mA TTS工作：10mA

方案二	POPEN口：低电平(接1K电阻到GND) MUTE口：悬空 (上位机IO口自行控制功放使能)	睡眠：不支持低功耗 启动：5mA 空闲：7mA TTS工作：10mA
方案三	POPEN口：高电平(默认，无需处理) MUTE口：接功放使能口	睡眠：1uA 启动：5mA 空闲：6mA TTS工作：9mA

【最低功耗模式推荐】

1、上位机通过电源管理芯片控制本芯片的供电脚“VCC”的通断电（例如：播放完毕断电，播放前上电）

注意：芯片VCC管脚断电后，要注意防止电流倒灌现象。

2、针对方案三，上位机通过串口发送睡眠指令“FD 00 01 22”让芯片进入睡眠状态(例如：播放完毕进入睡眠，播放前唤醒)。

注意：客户要注意防串口抖动，导致误唤醒芯片。

11.2 芯片启动和复位

注意：

芯片启动或复位后，大约N毫秒后UART口会返回0x4A回传（回传只要包含0x4A即可，例如：0x00 0x4A 或 0x00 0x34 0x4A 或 0xF0 0x4A），代表“初

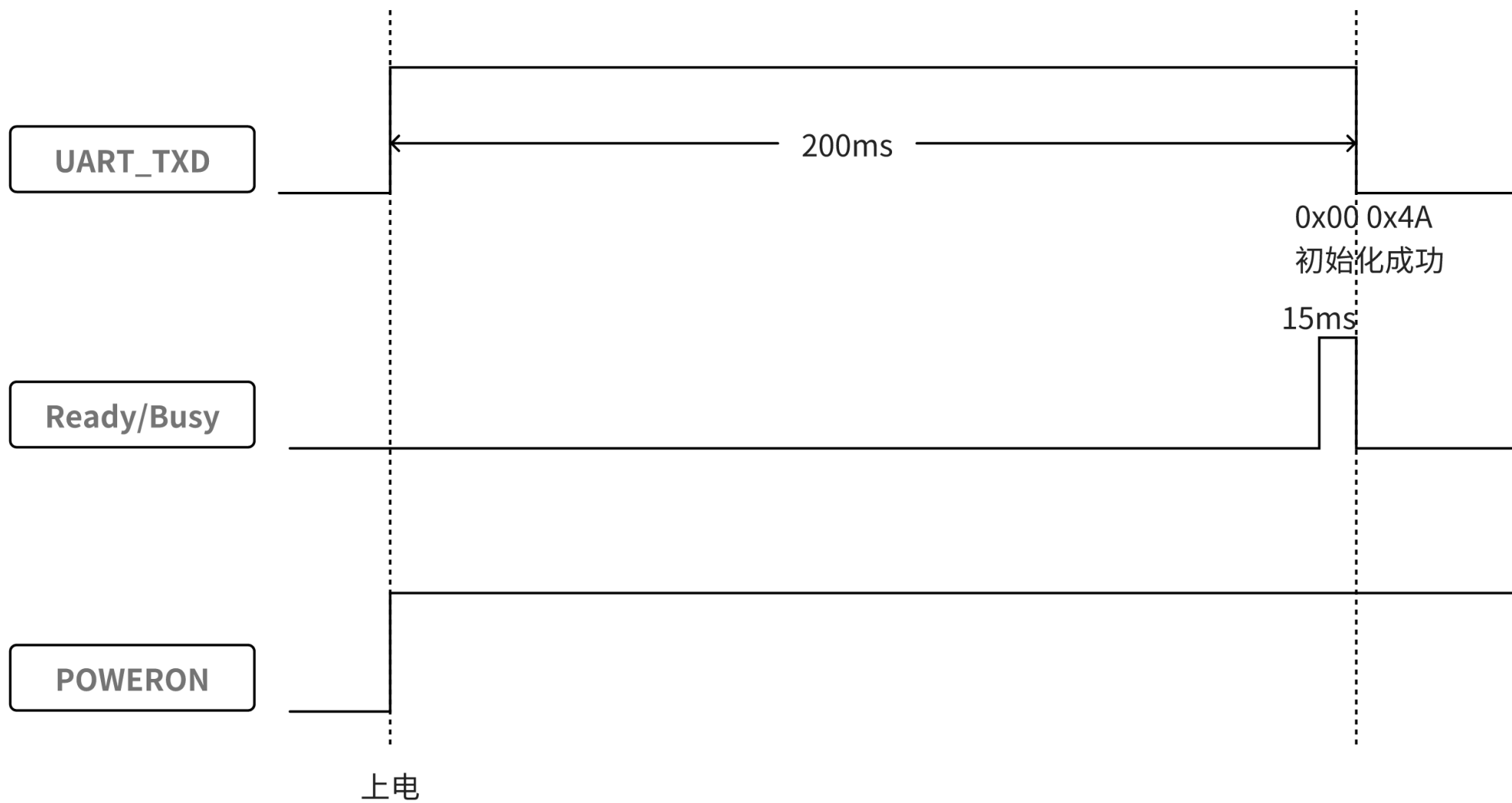
始化成功”，此时用户可以发送命令控制芯片。

注意：

上位机满足以下条件之一才能确认是“启动唤醒成功”：

- 1、收到0x4A回传；
- 2、启动过程中识别到“R/B脚”（即READY/BUSY脚）拉高15ms的信号；
- 3、不推荐用上电后固定N毫秒来算“启动唤醒成功”，因芯片的个体差异或者生产批次的不同会有不同变化。

本芯片启动时间约200ms左右（默认“去上电POP音延时=0”）

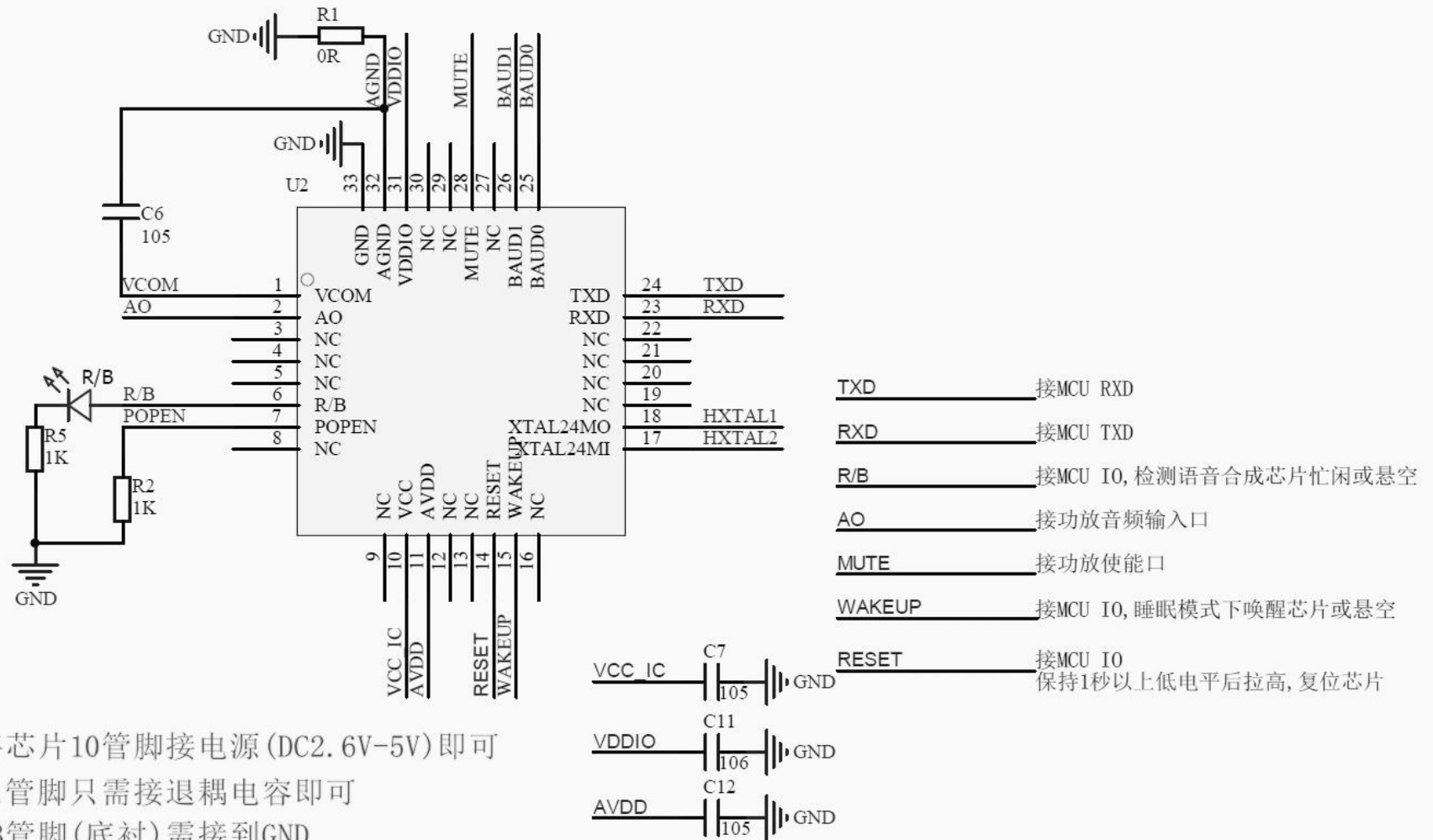


12. 芯片参考电路

12.1 免责声明

画板完成后仔细与我公司提供的参考电路进行比对，若客户未按照我公司提供的参考电路设计而造成的损失我司概不负责。如需对参考电路进行优化设计，请将优化后的电路图提交我们核对。

12.2 典型应用电路

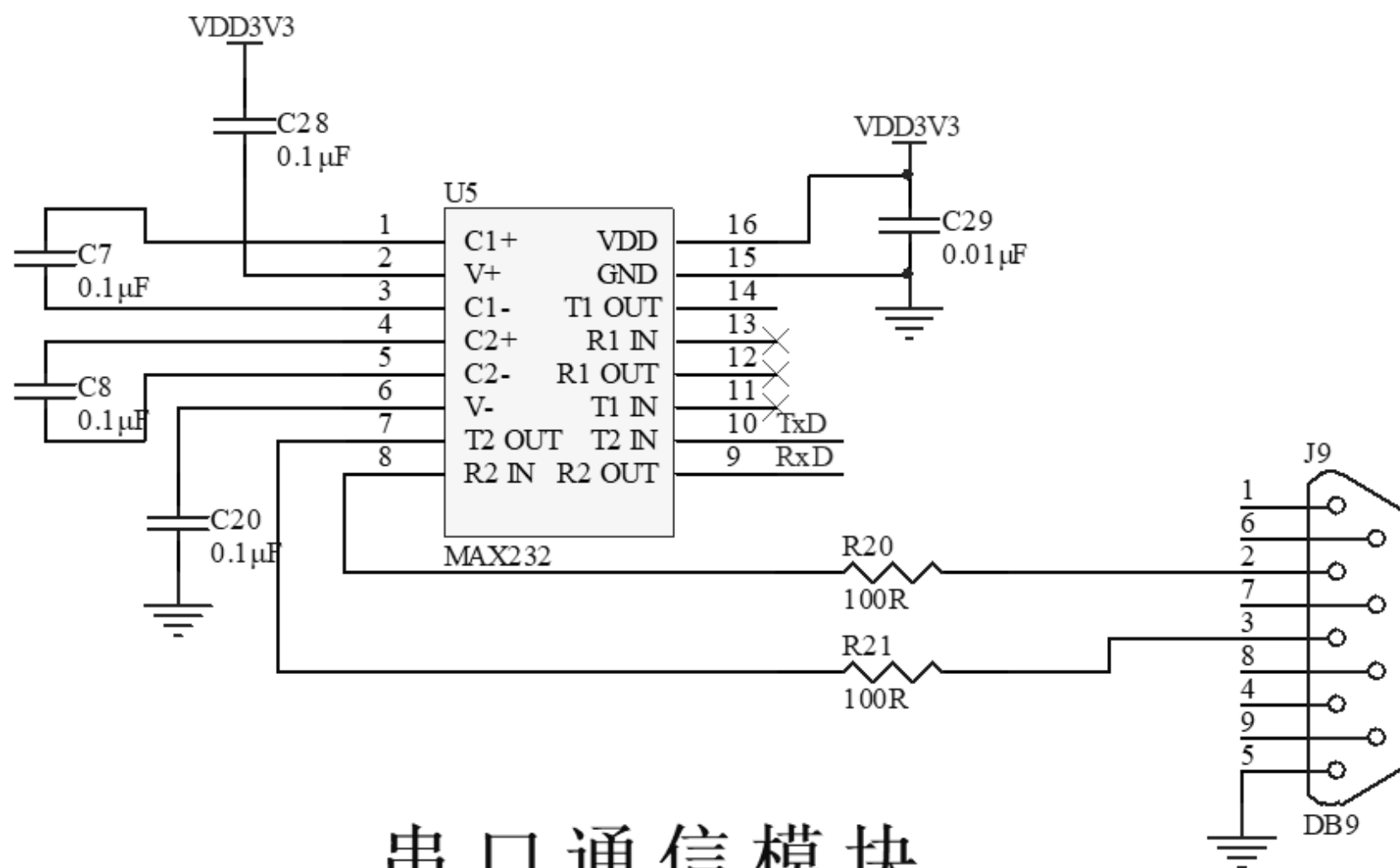


注意

只需将芯片10管脚接电源(DC2.6V-5V)即可
11和31管脚只需接退耦电容即可
芯片33管脚(底衬)需接到GND

12.3 上位机通讯连接

UART转串口与PC机通讯原理图:



串口通信模块

UART通讯

波特率配置方法以及R/B的应用：

本语音合成芯片的UART通信接口支持四种通信波特率：9600bps，57600bps，115200bps，460800bps，可以通过芯片上的两个管脚BAUD0和BAUD1 的电平来进行硬件配置.

0：表示低电平 1：表示高电平

波特率	BAUD0	BAUD1
9600 bps	0	1
57600 bps	1	0
115200 bps	1	1
460800 bps	0	0



注意：

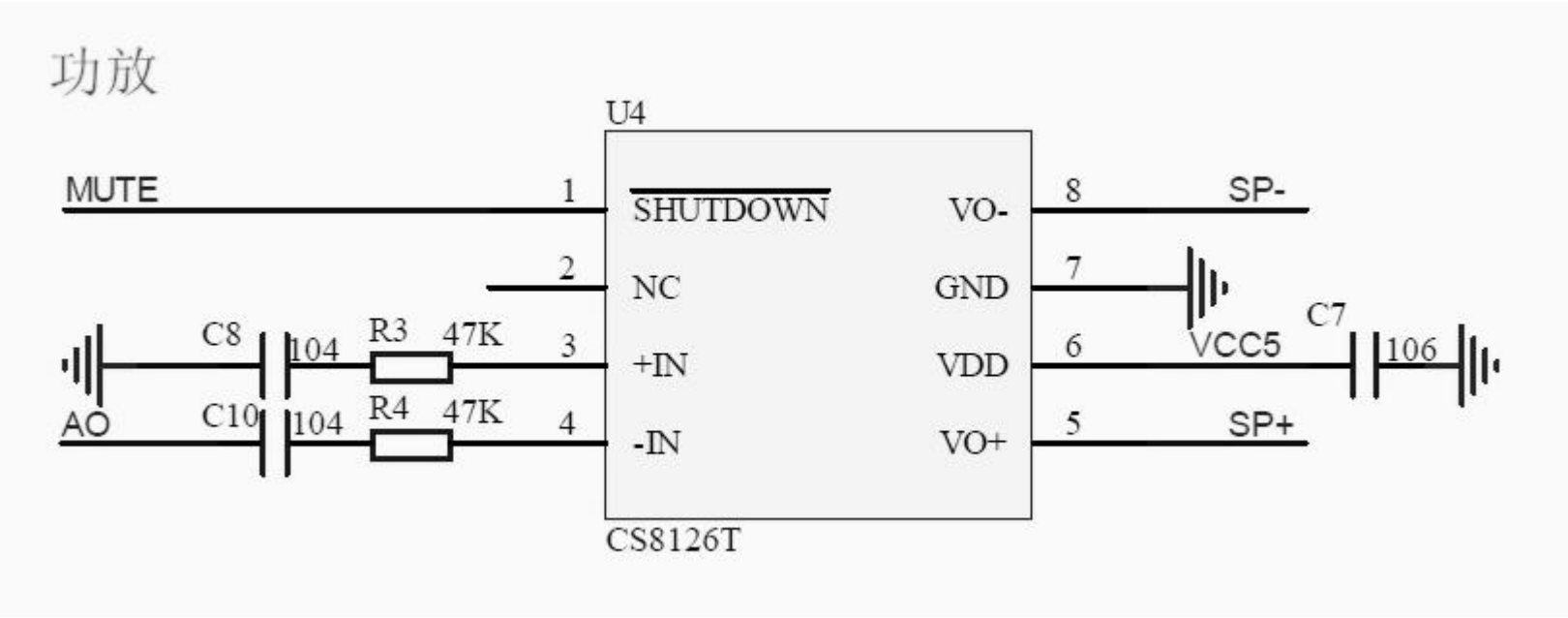
芯片通过管脚READY/BUSY的电平来指示芯片的工作状态，芯片空闲状态时READY/BUSY管脚输出低电平，芯片在语音合成状态时READY/BUSY管脚输出高电平。

12.4 电源电路

3.1V-5V单独供电，接VCC口。POPEN、BAUD0、BAUD1、WAKEUP默认均有上拉电阻，如果IO需要上拉请上拉到VDDIO，而不是VCC。

12.5 功放参考

CS8126T原理图：



CS8126T参数说明：

工作电压范围	VDD	RL	输出功率（典型值）	封装	备注

2.5V ~ 5.5V	3.6V	4Ω	1.3W	DFN8	-
	5V	4Ω	3.1W		



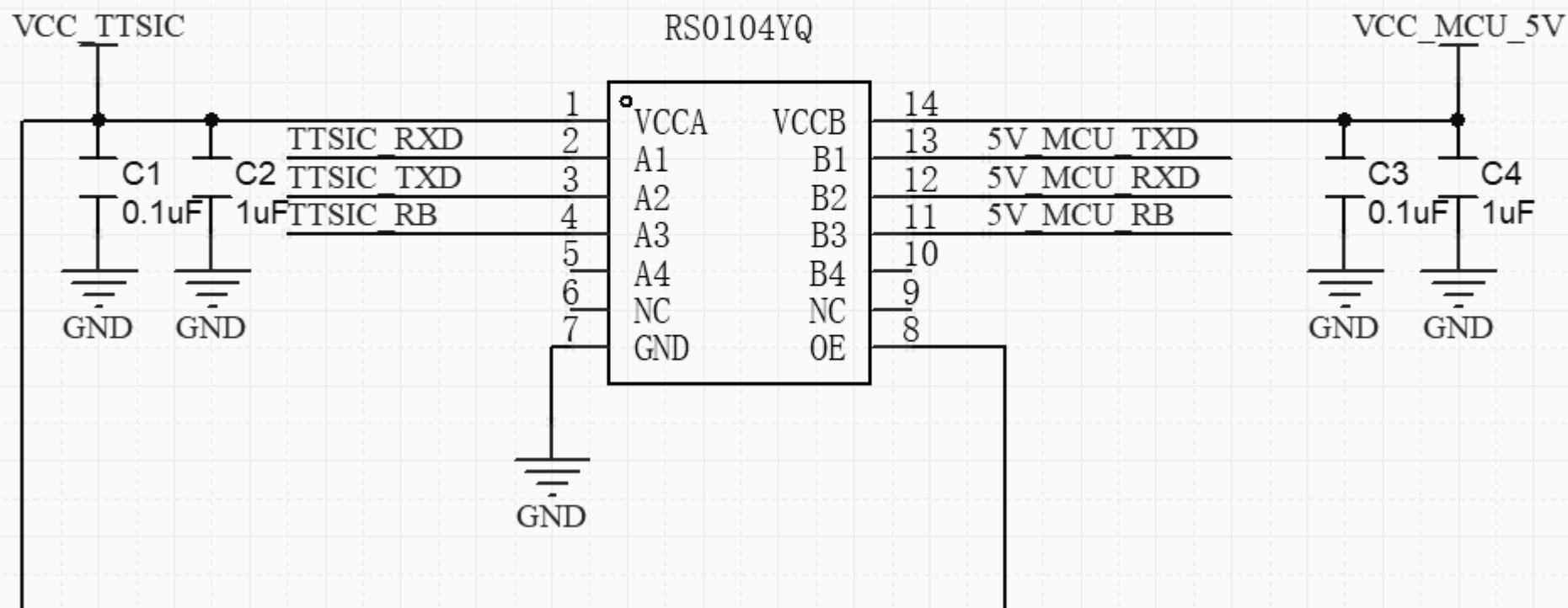
如需更多功放选型参考，请见 [《功放芯片推荐参考手册》文档](#)

12.6 5V MCU通信参考

5V MCU作为上位机与本芯片通信时，需要做电平转换，参考电路如下两种。推荐使用电平转换芯片电路，需要注意VCCB接的是电平较高的一端。

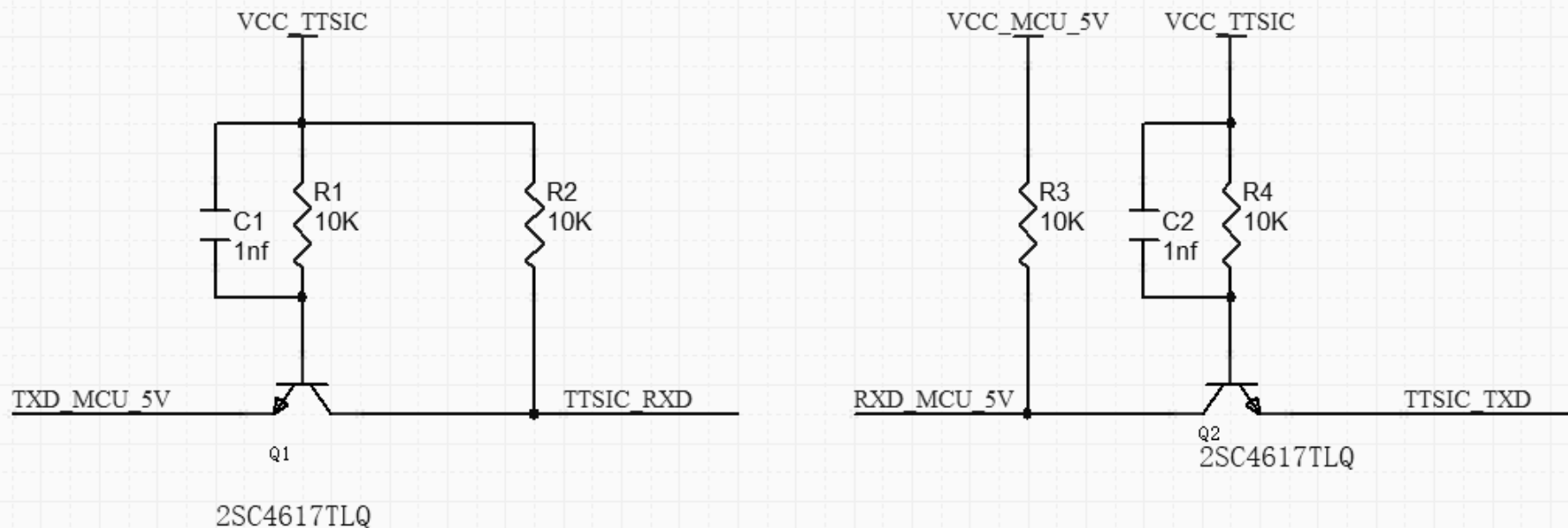
电平转换芯片参考电路

电平转换芯片参考电路



三级管电平信号转换参考电路

三极管电平信号转换参考电路



注意：

TTSIC代表本语音合成芯片。

13. 制板注意事项

13.1 电容布局

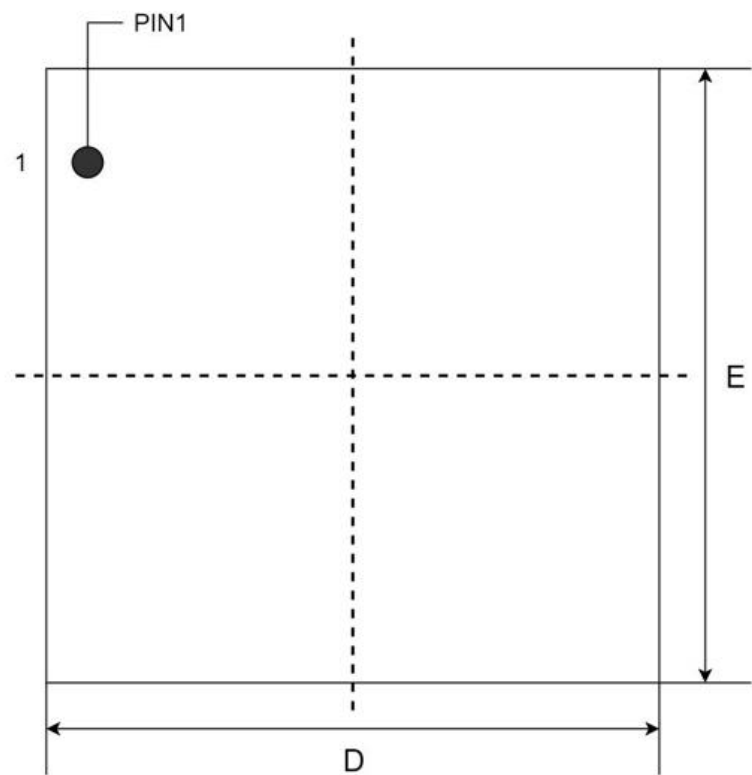
本芯片的各PIN口的对应滤波电容应尽量靠近其相应管脚。

13.2 音频线布线

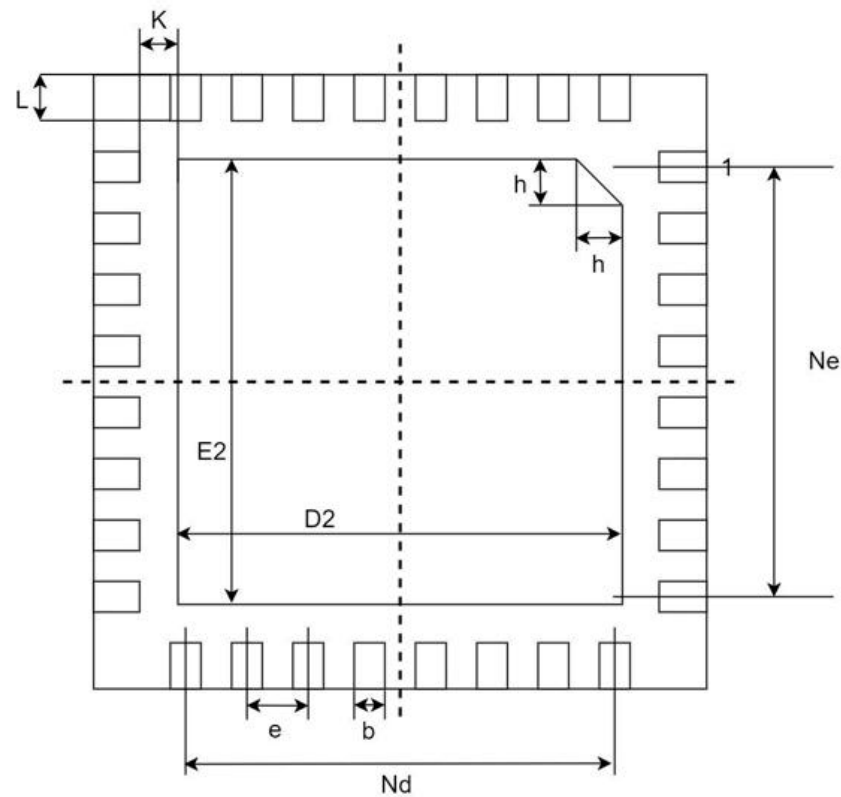
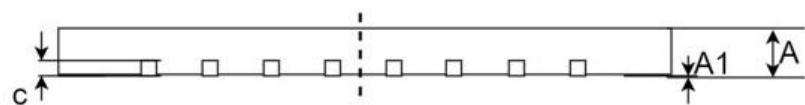
本芯片的AO，要包地、远离晶振及其他干扰源，否则可能会带来杂音。

14. 芯片规格

14.1 封装



TOP VIEW

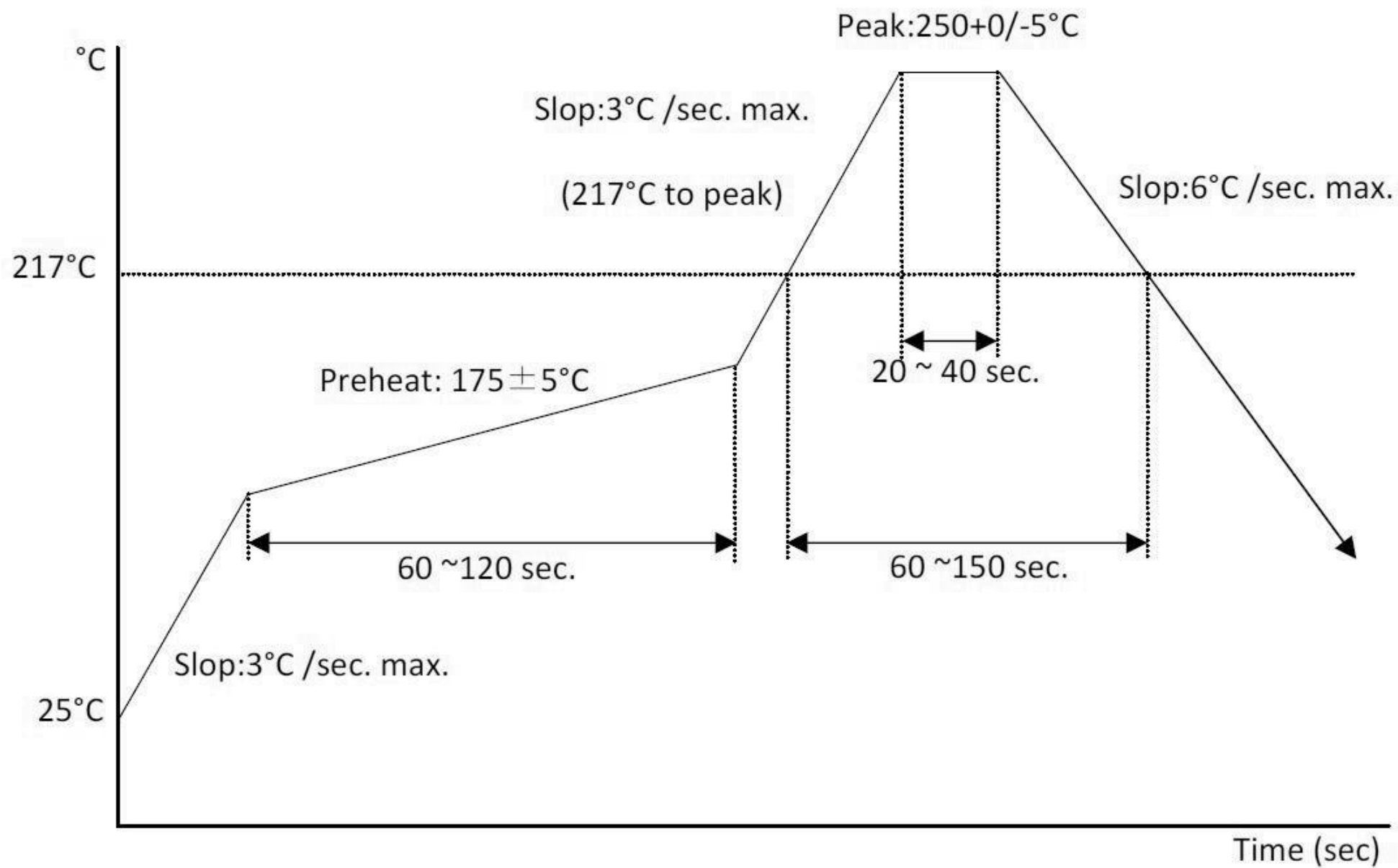


BOTTOM VIEW

标注	最小（mm）	标准（mm）	最大（mm）	标注	最小（mm）	标准（mm）	最大（mm）
A	0.7	0.8	0.9	Nd	2.80BSC		

A1	-	0.02	0.05		E	3.90	4.00	4.10
b	0.15	0.20	0.25		E2	2.80	2.90	3.00
e	0.40BSC				Ne	2.80BSC		
c	0.102REF				K	0.25REF		
D	3.90	4.00	4.10		L	0.25	0.30	0.35
D2	2.80	2.90	3.00		h	0.30	0.35	0.40

14.2 回流温度曲线



15. 芯片参数

15.1 特性极限值

参数	符号	最小值	最大值	单位
供电电压	VCC	2.6	5	V
3.3V IO输入电压	V _{3.3IO}	-0.3	VDDIO+0.3	V
工作温度	To	-40	85	°C
存储温度	Ts	-60	150	°C

15.2 电压工作范围

参数	符号	最小值	典型值	最大值	单位
供电电压	VCC	2.6	3.3	5	V

15.3 IO输入电气特性

参数	符号	最小值	典型值	最大值	单位	实验条件
低电平输入电压	V _{IL}	-0.3	-	0.3*VDDIO	V	VDDIO=3.3
高电平输入电压	V _{IH}	0.7*VDDIO	-	VDDIO+0.3	V	VDDIO=3.3

15.4 IO输出电气特性

参数	符号	最小值	典型值	最大值	单位	实验条件
低电平输出电压	V _{OL}	-	-	0.33	V	VDDIO=3.3
高电平输出电压	V _{OH}	2.7	-	-	V	VDDIO=3.3

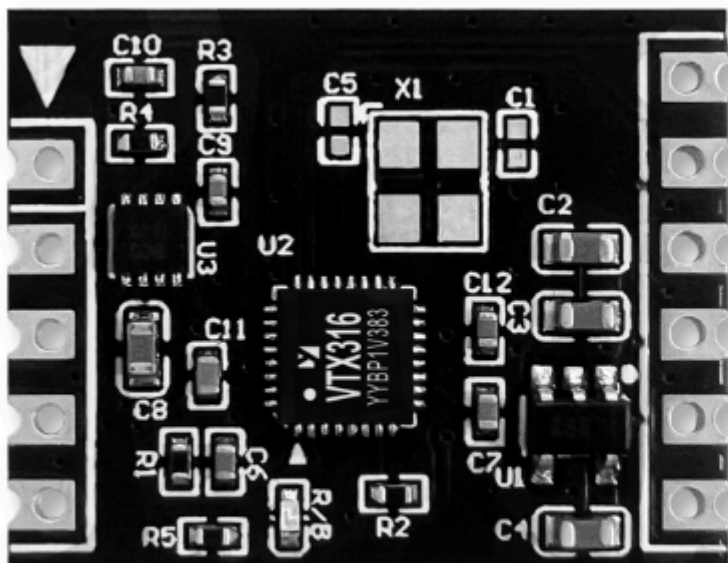
15.5 音频DAC特性

参数	最小值	典型值	最大值	单位	实验条件
频率响应	20	-	20K	Hz	1KHz/0dB 10Kohm负载，带A加权滤波器
THD+N	-	-75	-	dB	
S/N	-	95	-	dB	
串扰	-	-90	-	dB	
输出幅度	-	1	-	V _{rms}	
动态范围	-	90	-	dB	1KHz/-60dB 10Kohm负载，带加权滤波器
DAC输出功率	11	-	-	mW	32ohm负载

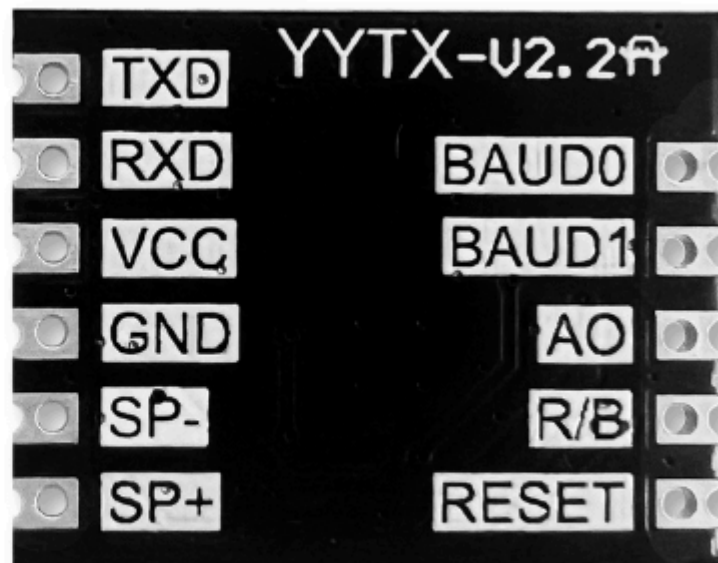
16. 模块概述

为方便客户评估测试语音合成芯片效果，我们制作成模块，客户只需要通过电源、UART、音频接口等几个简单的连线就能对语音合成芯片的性能做整体评估。

16.1 模块实物图



模块正视图

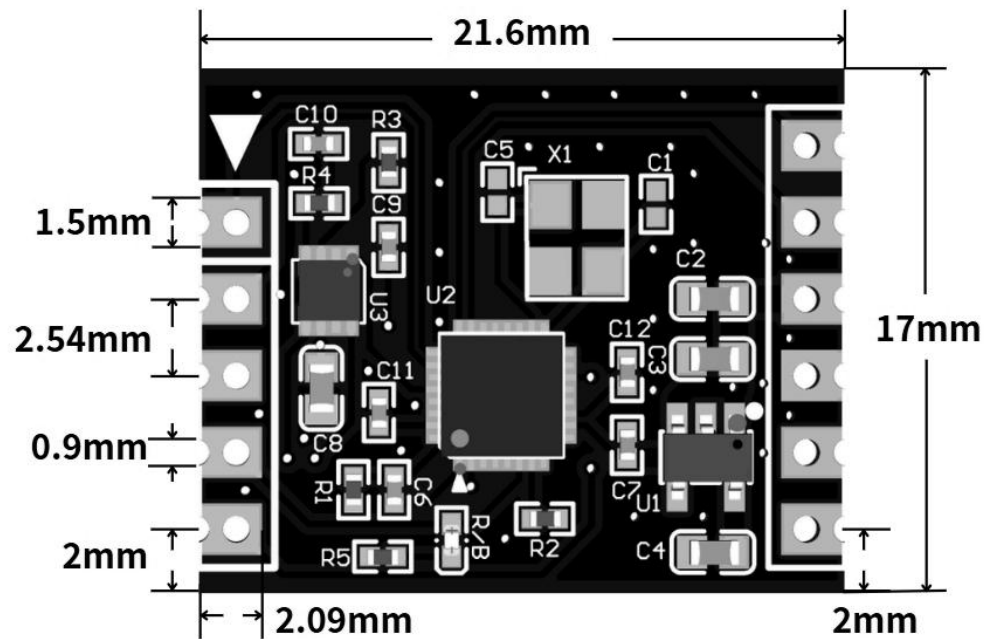


模块底视图

注意：

模块板载CS8126T功放芯片，详见[《功放芯片推荐参考手册》](#)

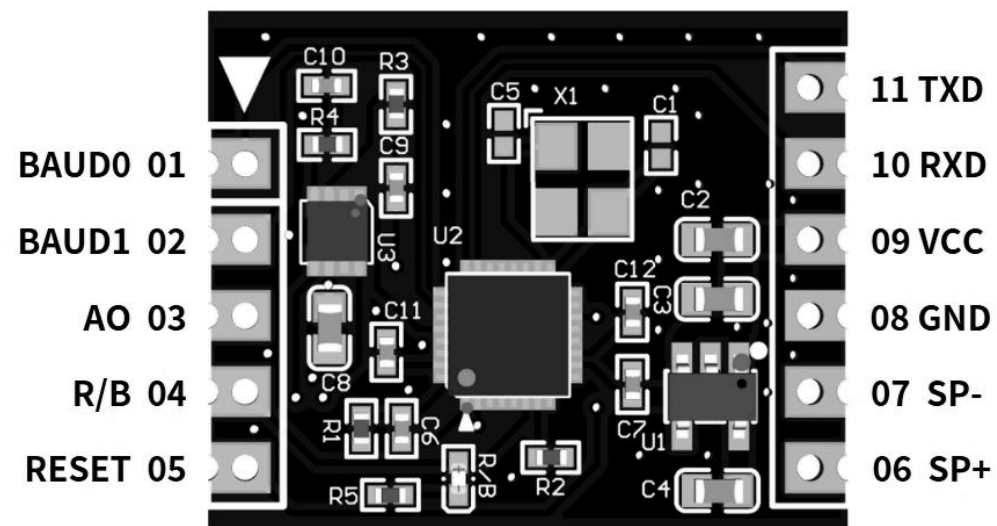
16.2 模块尺寸



注意：

模块厚度 $2.15 \pm 0.1\text{mm}$ (板厚 1.0mm)

16.3 模块引脚图



16.4 模块引脚定义

编号	引脚	说明	编号	引脚	说明
			11	TXD	模块串口发送
1	BAUD0	波特率配置端口0	10	RXD	模块串口接收
2	BAUD1	波特率配置端口1	9	VCC	5V供电
3	AO	音频输出	8	GND	GND

4	R/B	芯片忙状态指示，低电平表示空闲	7	SP-	接喇叭负
5	RESET	复位脚（保持1秒以上低电平复位）	6	SP+	接喇叭正

16.5 模块测试方法

测试工具准备：

- 电脑
- 喇叭
- USB转TTL串口线
- PC端演示程序

模块接线方法：

模块引脚与USB转TTL串口线、喇叭对应关系，见下表

模块引脚	USB转TTL串口线	喇叭
11 TXD	RXD	-
10 RXD	TXD	-
9 VCC	VCC（5V）	-

模块引脚	USB转TTL串口线	喇叭
8 GND	GND	-
7 SP-	-	SP-
6 SP+	-	SP+



注意：
请配套我司提供的PC端演示程序，程序在下方《资料下载》章节中

17. 资料下载

通用PC端演示程序

适用于宇音天下下位机语音芯片系列功能演示