



XFS5152CE 语音合成芯片

用户开发指南

合肥讯飞数码科技有限公司

安徽省合肥市望江西路 666 号 国家科技创新型试点市示范区科大讯飞语音产业基地



0551-65331506; 18256923366



0551-65331837



www.iflytek.com

版本历史

版 本	日 期	修改记录
V0.1	2013-4-15	初稿
V0.2	2013-4-23	修订电路参考图等细节
V0.3	2013-4-24	增加 12.1 节文本标记使用的详细方法说明，修订管脚定义等细节。
V1.0	2013-5-6	正式发布版本
V1.1	2013-5-13	修订 5 章表格中管脚定义标识，修订 8.3.4.1 表格、8.4.3 章节、8.4.4 章节、13.2 章节的相关内容。
V1.2	2013-5-14	修订文中提示音、识别命令词定制的相关说明。

声 明

本用户开发指南由合肥讯飞数码科技有限公司版权所有，未经许可，任何单位和个人都不得以电子的、机械的、磁性的、光学的、化学的、手工的等形式复制、传播、转录和保存该出版物，或翻译成其他语言版本。一经发现，将追究其法律责任。

讯飞数码保证本开发指南提供信息的准确性和可靠性，但并不对文本中可能出现的文字错误或疏漏负责。讯飞数码保留更改本开发指南的权利，如有修改，恕不相告。请在订购时联系我们以获得产品最新信息。对任何用户使用我们产品时侵犯第三方版权或其他权利的行为讯飞数码概不负责。另外，在讯飞数码未明确表示产品有该项用途时，对于产品使用在极端条件下导致一些失灵或损毁而造成的损失概不负责。

目 录

1	概述	- 1 -
2	主要应用领域	- 1 -
3	订货信息	- 1 -
4	功能描述	- 2 -
5	引脚定义	- 3 -
6	通讯接口介绍	- 5 -
6.1	UART 通讯模式	- 5 -
6.1.1	硬件连接	- 5 -
6.1.2	通讯传输字节格式	- 5 -
6.1.3	波特率配置方法	- 6 -
6.2	I2C 通讯模式	- 6 -
6.2.1	硬件连接	- 6 -
6.2.2	通讯传输字节格式	- 7 -
6.3	SPI 通讯模式	- 7 -
6.3.1	硬件连接	- 7 -
6.3.2	通讯传输字节格式	- 8 -
7	芯片控制方式	- 9 -
7.1	控制命令	- 9 -
7.2	芯片回传	- 9 -
8	通讯协议	- 11 -
8.1	查询和控制模式命令	- 11 -
8.1.1	状态查询命令	- 11 -
8.1.2	进入省电模式命令	- 11 -
8.1.3	唤醒命令	- 11 -
8.2	语音合成功能支持的控制命令	- 12 -
8.2.1	语音合成命令	- 13 -
8.2.2	停止合成命令	- 14 -
8.2.3	暂停合成命令	- 14 -
8.2.4	恢复合成命令	- 14 -
8.2.5	芯片状态回传	- 15 -
8.2.6	语音合成功能使用方法	- 15 -
8.2.6.1	简单调用方式	- 15 -
8.2.6.2	标准调用方式	- 15 -
8.2.6.3	查询芯片工作状态的方法	- 16 -
8.3	文本缓存功能支持的控制命令	- 16 -
8.3.1	命令帧定义	- 16 -

8.3.2	发送缓存文本命令.....	- 17 -
8.3.3	播放缓存文本命令.....	- 17 -
8.3.4	发送及播放缓存文本的举例.....	- 18 -
8.3.4.1	发送缓存文本命令.....	- 18 -
8.3.4.2	文本缓存播放命令.....	- 19 -
8.4	语音编解码功能支持的控制命令.....	- 19 -
8.4.1	命令帧定义.....	- 19 -
8.4.2	开始语音编码的命令.....	- 20 -
8.4.3	开始语音解码的命令.....	- 20 -
8.4.4	发送数据到芯片进行解码.....	- 22 -
8.4.5	停止语音编解码的命令.....	- 22 -
8.4.6	编解码相关的状态回传.....	- 22 -
8.4.7	编解码延迟时间参数.....	- 23 -
8.5	语音识别功能支持的控制命令.....	- 23 -
8.5.1	命令帧定义.....	- 23 -
8.5.1	启动语音识别命令.....	- 23 -
8.5.2	停止语音识别命令.....	- 24 -
8.5.3	芯片语音识别结果回传.....	- 24 -
9	语音芯片系统结构图.....	- 25 -
10	参考电路.....	- 26 -
10.1	典型应用电路.....	- 26 -
10.2	电源模块参考电路.....	- 26 -
10.3	晶振电路.....	- 27 -
10.4	复位电路.....	- 27 -
10.5	与上位机通讯的电路连接.....	- 27 -
10.6	录音模块电路.....	- 28 -
10.7	音频输出功放电路.....	- 28 -
10.7.1	TPA3123D2.....	- 29 -
10.7.2	TPA3110D1.....	- 29 -
10.7.3	TPA2005D1.....	- 29 -
11	示例程序.....	- 31 -
11.1	发送语音合成文本的示例程序.....	- 31 -
11.2	语音编解码的示例程序.....	- 32 -
12	文本控制标记.....	- 37 -
12.1	文本控制标记的使用方法.....	- 37 -
12.2	文本控制标记列表.....	- 37 -
12.3	使用示例.....	- 39 -
12.3.1	[f?]标记的使用.....	- 39 -
12.3.2	[g?] 标记的使用.....	- 39 -

12.3.3	[h?] 标记的使用	- 39 -
12.3.4	[i?]标记的使用	- 39 -
12.3.5	[m?]标记的使用	- 40 -
12.3.6	[n?]标记的使用	- 40 -
12.3.7	[o?]标记的使用	- 40 -
12.3.8	[p?]标记的使用	- 40 -
12.3.9	[r?]标记的使用	- 41 -
12.3.10	[s?]标记的使用	- 41 -
12.3.11	[t?]标记的使用	- 41 -
12.3.12	[v?]标记的使用	- 41 -
12.3.13	[x?]标记的使用	- 42 -
12.3.14	[y?]标记的使用	- 42 -
12.3.15	[z?]标记的使用	- 42 -
12.3.16	[=?]标记的使用	- 42 -
12.3.17	[d]标记的使用	- 42 -
13	提示音	- 44 -
13.1	芯片集成提示音	- 44 -
13.1.1	信息提示音列表	- 44 -
13.1.2	铃声提示音列表	- 44 -
13.1.3	警报提示音列表	- 44 -
13.2	提示音定制功能	- 45 -
14	语音识别命令词	- 46 -
14.1	命令词列表	- 46 -
14.2	命令词定制说明	- 46 -
15	产品规格	- 47 -
15.1	封装	- 47 -
15.2	特性参数	- 48 -
15.2.1	极限值	- 48 -
15.2.2	直流电气特性	- 48 -
15.2.3	音频 DAC 特性	- 48 -
15.2.4	非文本合成状态下的功耗参数	- 49 -
15.3	接收合成命令到开始播音间隔时间	- 49 -
15.4	焊接工艺要求	- 49 -
15.4.1	烘烤温度及时间	- 49 -
15.4.2	回流焊的峰值温度	- 50 -
16	附录	- 51 -
16.1	GB2312 编码简介	- 51 -
16.2	UNICODE 编码简介	- 51 -
16.3	GBK 编码简介	- 51 -

16.4	BIG5 编码简介	- 52 -
------	-----------------	--------

1 概述

XFS5152CE 是一款高集成度的语音合成芯片，可实现中文、英文语音合成；并集成了语音编码、解码功能，可支持用户进行录音和播放；除此之外，还创新性地集成了轻量级的语音识别功能，支持 30 个命令词的识别，并且支持用户的命令词定制需求。

2 主要应用领域

- 车载调度终端
- 信息机
- 考勤机
- 排队机
- 气象预警机
- 智能仪器
- 固定电话
- 税控机
- 公交车语音报站器
- 自动售货机
- POS 机
- 智能仪表

3 订货信息

芯片型号	封装信息	
	封装名称	封装描述
XFS5152CE	LQFP64	64 脚，芯片 10mmx10mmx1.4mm

表格 1 订货信息

4 功能描述

- 支持任意中文文本、英文文本的合成，并且支持中英文混读

芯片支持任意中文、英文文本的合成，可以采用 GB2312、GBK、BIG5 和 UNICODE 四种编码方式。每次合成的文本量最多可达 4K 字节。

芯片对文本进行分析，对常见的数字、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；对一般多音字也可以依据其语境正确判断读法；另外针对同时有中文和英文的文本，可实现中英文混读。

- 支持语音编解码功能，用户可以使用芯片直接进行录音和播放

芯片内部集成了语音编码单元和解码单元，可以进行语音的编码和解码，实现录音和播放功能。芯片的语音编解码具备高压缩率、低失真率、低延时的特点，并且可以支持多种语音编解码速率。这些特性使它非常适合于数字语音通信、语音存储以及其它需要对语音进行数字处理的场合。如：车载微信、指挥中心等。

- 支持语音识别功能

可支持 30 个命令词的识别。芯片出默认设置的是 30 个车载、预警等行业常用识别命令词。客户如需要更改成其他的识别命令词，可进行命令词定制。

- 芯片内部集成 80 种常用提示音效

适合用于不同场合的信息提示、铃声、警报等功能。

- 支持 UART、I2C 、SPI 三种通讯方式

UART 串口支持 4 种通讯波特率可设：4800 bps、9600 bps、57600 bps、115200 bps,用户可以根据情况通过硬件配置选择自己所需的波特率。

- 支持多种控制命令

如合成文本、停止合成、暂停合成、恢复合成、状态查询、进入省电模式、唤醒等。

控制器通过通讯接口发送控制命令可以对芯片进行相应的控制。芯片的控制命令非常简单易用，例如：芯片可通过统一的“合成命令”接口播放提示音和中文文本，还可以通过标记文本实现对合成的参数设置。

- 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工作状态字、发送查询命令获得芯片工作状态的回传数据。

5 引脚定义

编号	引脚	说明	编号	引脚	说明
1	VDD12	1.2V 数字电源输入	33	NC	——
2	UVCC	3.3V 模拟电源输入	34	NC	——
3	RREF	参考电压，接 10K 下拉电阻	35	NC	——
4	GND	数字地	36	VDD33	3.3V 数字电源输入
5	NC	——	37	GND	数字地
6	NC	——	38	NC	——
7	AVCC	3.3V 模拟电源输入	39	RXD	串口接收
8	AO_P	音频输出正	40	RDY	低电平芯片处于就绪状态
9	AO_N	音频输出负	41	TXD	串口发送
10	AGND	模拟地	42	SPI_CLK	SPI 时钟输入
11	VCM	模拟参考电压	43	SPI_SSEL	SPI 片选，接 10K 上拉电阻
12	Mic_Bias	麦克偏置电压	44	SPI_SI	SPI 数据输入
13	MIC_N	麦克输入负，不用时悬空	45	SPI_SO	SPI 数据输出
14	AGND	模拟地	46	VDD33	3.3V 数字电源输入
15	MIC_P	麦克输入正，不用时悬空	47	IIC_SCL	IIC 时钟输入
16	F_D0	信号脚，与 57 脚相连	48	IIC_SDA	IIC 数据接口
17	AVCC	3.3V 模拟电源输入	49	M_D2	信号脚，与 50 脚相连，加 10K 上拉
18	VREF	参考电压，外接 0.1uF 电容到地	50	F_D2	信号脚，与 49 脚相连，加 10K 上拉
19	NC	——	51	F_D3	信号脚，与 59 脚相连
20	PWR_IN	芯片供电，3.3V	52	GND	数字地
21	F_D1	信号脚，与 60 脚相连	53	A1	与 54 脚相连
22	VDD33_O	3.3V 电源输出	54	A2	与 53 脚相连
23	VDD12_O	1.2V 电源输出	55	BAUD2	波特率选择引脚

24	AVDD12	1.2V 模拟电源输出	56	BAUD1	波特率选择引脚
25	GND	数字地	57	M_D0	信号脚, 与 16 脚相连
26	XTALI	晶振输入	58	NC	——
27	XTALO	晶振输出	59	M_D3	信号脚, 与 51 脚相连
28	RESET	复位引脚, 低有效	60	M_D1	信号脚, 与 21 脚相连
29	A4	外加 10K 上拉电阻	61	NC	——
30	NC	——	62	A3	外加 10K 下拉电阻
31	NC	——	63	NC	——
32	NC	——	64	GND	数字地

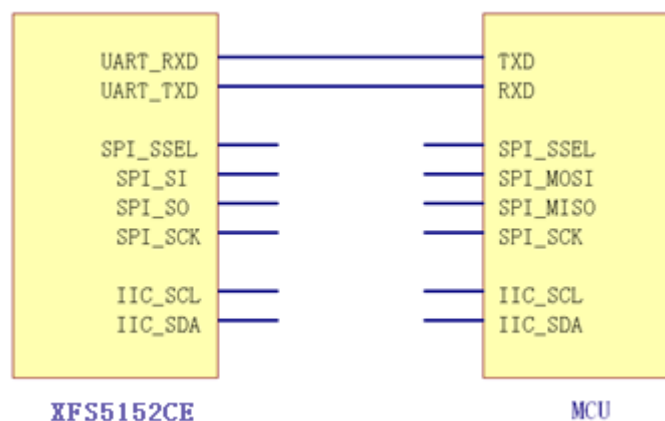
表格 2 XFS5152CE 的引脚定义表

6 通讯接口介绍

XFS5152CE 芯片支持 UART 接口、I2C 接口、SPI 接口三种通讯方式，可通过 UART 接口、I2C 或 SPI 接口接收上位机发送的命令和数据，允许发送数据的最大长度为 4k 字节。

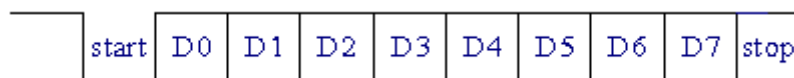
6.1 UART 通讯模式

6.1.1 硬件连接



图表 1 XFS5152CE 芯片 UART 接口的连接示意图

6.1.2 通讯传输字节格式



图表 2 UART 接口通讯传输字节格式

- 1) 通讯标准：UART
- 2) 波特率：4800 bps、9600 bps、57600 bps、115200 bps
- 3) 起始位：1bit
- 4) 数据位：8 bits
- 5) 停止位：1 bit
- 6) 校验：无

6.1.3 波特率配置方法

XFS5152CE 芯片的 UART 通讯接口支持 4 种通讯波特率：4800 bps、9600 bps、57600 bps、115200 bps，可以通过 XFS5152CE 芯片上的两个管脚 BAUD1（56 引脚）、BAUD2（55 引脚）上的电平来进行硬件配置。

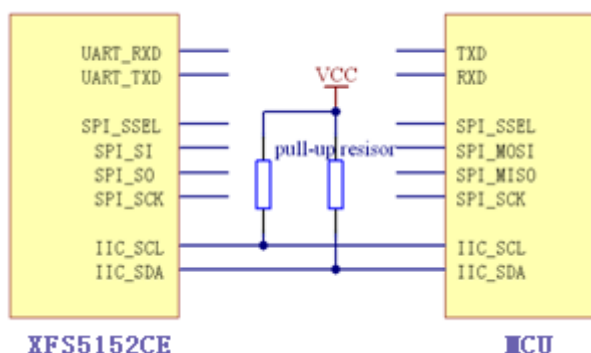
波特率	BAUD1	BAUD2
4800 bps	0	0
9600 bps	0	1
57600 bps	1	0
115200 bps	1	1

表格 3 波特率配置表

6.2 I2C 通讯模式

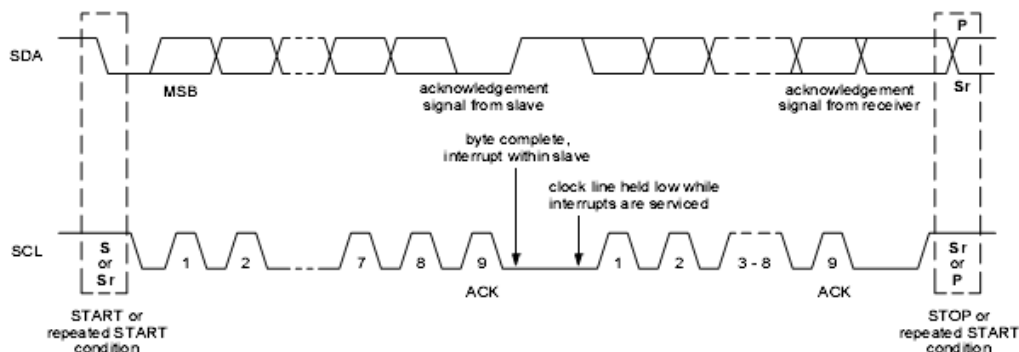
6.2.1 硬件连接

XFS5152CE 采用标准 I²C 总线接口，设置为 Slave 身份。I²C 总线是一个为简化多个智能设备之间相互传输信息而设计的两线、双向网络。



图表 3 XFS5152CE 芯片 I2C 接口的连接示意图

6.2.2 通讯传输字节格式



图表 4 I²C 总线时序图

XFS5152CE 作为 I2C 接口的从设备，设备地址为：**0x80**。

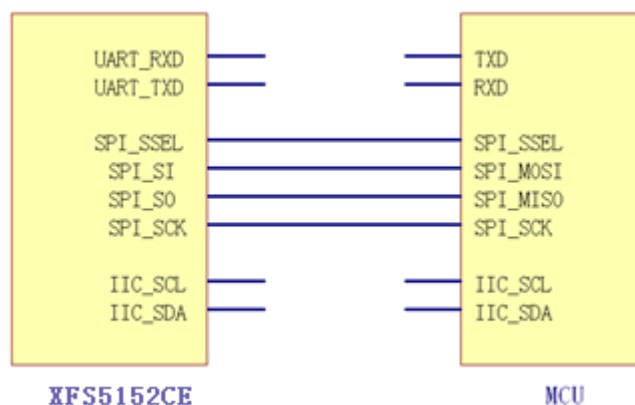
为保证通讯的稳定性，XFS5152CE 的 I2C 接口通信速度要求不高于 15KHZ。

XFS5152CE 采用的标准 I²C 总线接口，关于 I²C 总线通讯的详细信息，请参考“The I²C-bus specification”。

6.3 SPI 通讯模式

6.3.1 硬件连接

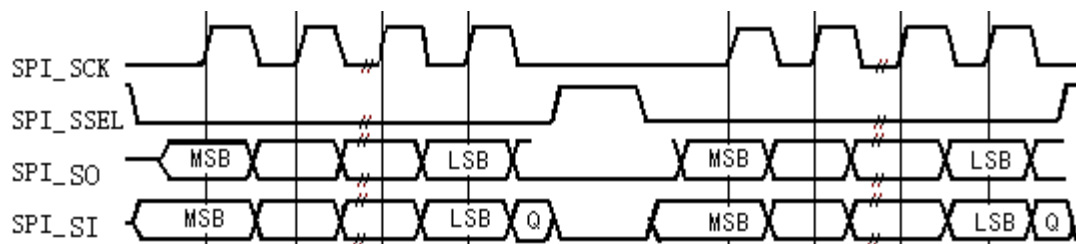
XFS5152CE 芯片的 SPI 接口是 4 线全双工同步串行通讯接口，上位机作为 SPI 通讯中的 Master 身份，XFS5152CE 在 SPI 通讯中设为 Slave 身份，SPI 通讯所需的时钟信号由上位机提供。



图表 5 XFS5152CE 芯片 SPI 接口的连接示意图

6.3.2 通讯传输字节格式

SPI 通讯属于同步串行通讯，为保证通讯的稳定性，XFS5152CE 的 SPI 接口通信速度要求不高于 25KHZ。上位机在和 XFS5152CE 的通讯中，由上位机提供同步时钟信号，在同步时钟的上升沿 XFS5152CE 锁存 1bit 数据，每传输 8bits 数据完成一个字节数据的传输。下图是 XFS5152CE 的 SPI 接口的通讯时序图。



图表 6 SPI 通讯时序图

注意：在发送一个字节之后，将 SPI_SCK 保持低电平一段时间，每两个字节之间必须插入一段等待，时间不少于 105 us。

7 芯片控制方式

7.1 控制命令

XFS5152CE 支持多种控制命令，列表如下：

命令功能		说明
语音合成相关	合成命令	合成本次发送的命令
	停止合成命令	停止当前的合成动作
	暂停合成命令	暂停正在进行的合成
	恢复合成命令	继续合成被暂停的文本
文本缓存相关	发送缓存文本命令	把需要缓存的文本发送到芯片缓存区特定的区段
	播放缓存文本命令	播放缓存区内已存储的多段文本
语音编解码相关	开始语音编码命令	设置语音编码的参数，并启动语音编码
	开始语音解码命令	设置语音解码的参数，并启动语音解码
	发送数据到芯片进行解码命令	根据指定的压缩等级发送一帧数据给芯片进行解码播放
	停止语音编解码命令	停止执行语音编码或解码操作
语音识别相关	启动语音识别命令	启动语音识别功能
	停止语音识别命令	停止当前的识别
	状态查询命令	查询当前芯片的工作状态
	进入省电模式	使芯片从正常工作模式进入省电模式
	唤醒命令	使芯片从省电模式进入正常工作模式

表格 4 控制命令列表

7.2 芯片回传

上电初始化成功时或者收到上位机的控制命令帧时，XFS5152CE 芯片都会向上位机发送 1 个字节的回传，上位机可根据这个回传来判断芯片目前的状态。

名称	回传信息	触发条件
初始化成功回传	0x4A	上电芯片初始化成功后，芯片自动发送回传
收到正确的命令帧回传	0x41	收到正确的命令帧

收到不能识别命令帧回传	0x45	收到错误的命令帧
芯片忙碌状态回传	0x4E	芯片处在正在合成状态，收到状态查询命令帧
芯片空闲状态回传	0x4F	芯片处在空闲状态，收到状态查询命令帧； 或一帧数据合成结束，芯片处于空闲状态

表格 5 芯片回传列表

8 通讯协议

8.1 查询和控制模式命令

8.1.1 状态查询命令

名称	发送的数据	说明		
命令字	0x21	通过该命令获取相应参数，来判断 TTS 芯片是否处在合成状态，返回 0x4E 表明芯片仍在合成中，返回 0x4F 表明芯片处于空闲状态		
参数列表	无			
命令帧格式结构				
	帧头	数据区长度		数据区
	0xFD	高字节	低字节	命令字
		0x00	0x01	0x21

表格 6 状态查询命令

8.1.2 进入省电模式命令

名称	发送的数据	说明		
命令字	0x88	进入省电模式		
参数列表	无			
命令帧格式结构				
	帧头	数据区长度		数据区
	0xFD	高字节	低字节	命令字
		0x00	0x01	0x88

表格 7 进入省电模式命令

8.1.3 唤醒命令

名称	发送的数据	说明
命令字	0xFF	从省电模式返回正常工作模式
参数列表	无	

命令帧格式结构	帧头	数据区长度		数据区
	0xFD	高字节	低字节	命令字
		0x00	0x01	0xFF

表格 8 唤醒命令

注意：芯片进入省电模式后,需要发送唤醒命令到芯片的通讯接口上,把芯片从省电模式中唤醒,进入正常的工作模式（待合成状态），才可以进行合成操作。

8.2 语音合成功能支持的控制命令

上位机发送给 XFS5152CE 芯片的所有语音合成命令都需要用“帧”的方式进行封装后传输。帧结构由帧头标志、数据区长度和数据区三部分组成。

帧头	数据区长度	数据区
0xFD	0xFF, 0xFF	Data

表格 9 命令帧封装格式

名称	长度	说明
帧头	1 Byte	定义为十六进制“0xFD”
数据区长度	2 Bytes	用两个字节表示，高字节在前，低字节在后
数据区	小于 4k Bytes	命令字和命令参数，长度和“数据区长度”一致

表格 10 命令帧说明

数据区是由命令字和命令参数组成的，上位机使用命令字来实现语音合成芯片的各种功能。

名称	发送的数据	说明
命令字	0x01	语音合成命令
	0x02	停止合成命令，没有参数
	0x03	暂停合成命令，没有参数
	0x04	恢复合成命令，没有参数
	0x21	芯片状态查询命令
	0x88	芯片进入省电模式
	0xFF	芯片从省电模式返回正常工作模式
命令参数	不同命令字有不同参数列表，详见各命令字说明	

表格 11 数据区命令字和命令参数

8.2.1 语音合成命令

名称	发送的数据	说明							
命令字	0x01	带文本编码设置的文本播放命令							
参数列表	0Xxx	1Byte 表示文本的编码格式，取值为 0~3	参数取值	文本编码格式					
			0x00	GB2312					
			0x01	GBK					
			0x02	BIG5					
			0x03	UNICODE					
Data	待合成文本的二进制内容								
命令帧格式结构	帧头	数据区长度		数据区					
	0xFD	高字节	低字节	命令字	文本编码格式	待合成文本			
		0xHH	0xLL	0x01	0x00~0x03			
示例	语音合成命令帧示例，发送文本编码格式为“GB2312”的文本“科大讯飞”								
	0xFD	0x00	0x0A	0x01	0x00	0xBF	0xC6	0xB4	0xF3
	0xD1	0xB6	0xB7	0xC9					
	语音合成命令帧示例，发送文本编码格式为“GBK”的文本“科大讯飞”								
	0xFD	0x00	0x0A	0x01	0x01	0xBF	0xC6	0xB4	0xF3
	0xD3	0x8D	0xEF	0x77					
	语音合成命令帧示例，发送文本编码格式为“BIG5”的文本“科大讯飞”								
	0xFD	0x00	0x0A	0x01	0x02	0xAC	0xEC	0xA4	0x6A
	0xB0	0x54	0xAD	0xB8					
	语音合成命令帧示例，发送文本编码格式为“UNICODE”的文本“科大讯飞”								
	0xFD	0x00	0x0A	0x01	0x03	0xD1	0x79	0x27	0x59
	0xAF	0x8B	0xDE	0x98					
	特别说明	当 XFS5152CE 芯片正在合成文本的时候，如果又接收到一帧有效的合成命令帧，芯片会立即停止当前正在合成的文本，转而合成新收到的文本。							

表格 12 语音合成命令

8.2.2 停止合成命令

名称	发送的数据	说明		
命令字	0x02	停止当前合成		
参数列表	无			
命令帧格式结构				
	帧头	数据区长度		数据区
	0xFD	高字节	低字节	命令字
		0x00	0x01	0x02

表格 13 停止合成命令

8.2.3 暂停合成命令

名称	发送的数据	说明		
命令字	0x03	暂停当前合成		
参数列表	无			
命令帧格式结构				
	帧头	数据区长度		数据区
	0xFD	高字节	低字节	命令字
		0x00	0x01	0x03

表格 14 暂停合成命令

8.2.4 恢复合成命令

名称	发送的数据	说明		
命令字	0x04	恢复暂停的合成		
参数列表	无			
命令帧格式结构				
	帧头	数据区长度		数据区
	0xFD	高字节	低字节	命令字
		0x00	0x01	0x04

表格 15 恢复合成命令

8.2.5 芯片状态回传

XFS5152CE 芯片在上电初始化成功时会向上位机发送一个字节的“初始化成功”回传，初始化不成功时不发送此回传。

XFS5152CE 芯片收到一个命令帧后会判断此命令帧正确与否，如果命令帧正确返回“收到正确命令帧”回传，如果命令帧错误则返回“收到错误命令帧”回传。

XFS5152CE 芯片收到状态查询命令时，如果芯片正处于合成状态则返回“芯片忙碌”回传，如果芯片处于空闲状态则返回“芯片空闲”回传。

XFS5152CE 芯片在一帧数据合成完毕后，会自动返回一次“芯片空闲”的回传。

回传数据类型	回传数据	触发条件
初始化成功	0x4A	芯片初始化成功
收到正确命令帧	0x41	收到正确的命令帧
收到错误命令帧	0x45	收到错误的命令帧
芯片忙碌	0x4E	收到“状态查询命令”，芯片处于合成文本状态回传 0x4E
芯片空闲	0x4F	当一帧数据合成完以后，芯片进入空闲状态回传 0x4F； 当芯片收到“状态查询命令”，芯片处于空闲状态回传 0x4F

表格 16 芯片回传

8.2.6 语音合成功能使用方法

8.2.6.1 简单调用方式

简单调用是指用户不用关心 XFS5152CE 当前的工作状态，只需要发送文本到 XFS5152CE 合成为语音输出。

在简单调用情况下，上位机只要与 XFS5152CE 之间建立起 UART、SPI、或 I2C 通信方式中的一种，即可发送合成命令来实现文本的合成，上位机不需要判断 XFS5152CE 的回传数据或状态引脚的输出。

注意：如前一帧文本还没有合成完，就再次发送文本到 XFS5152CE 会中断前次合成，而执行新的合成。

8.2.6.2 标准调用方式

若上位机需要确保上次文本被完整合成之后，再发送合成命令帧合成下一段文本，则需要通过回传确定芯片的工作状态。具体方法举例如下：应用中需要合成的文本为 5k 字节，超过了

XFS5152CE 芯片一个命令帧所能容纳的最大文本长度(4k 字节),这时需要分两次给 XFS5152CE 芯片发送文本信息。程序过程如下:

- 1、上位机先给 XFS5152CE 芯片发送一个文本合成命令帧,携带不超过 4k 个字节的文本;
- 2、上位机等待 XFS5152CE 芯片自动返回的回传信息,直到收到“0x4F”回传,说明前面的文本已合成完毕;或使用查询芯片的状态引脚、发送查询命令的方法,确认上一帧文本合成完毕。
- 3、上位机向芯片再次发送一个文本合成命令帧,发送出剩下的文本。

8.2.6.3 查询芯片工作状态的方法

通过硬件和软件两种方式查询 XFS5152CE 芯片的工作状态。

硬件方式:通过查询输出引脚 \overline{RDY} 的电平,来判断芯片的工作状态。当 \overline{RDY} 处于高电平时,表明芯片正在合成文本;当 \overline{RDY} 处于低电平时,表明芯片处于空闲状态。

软件方式:通过发送状态查询命令帧来查询芯片的工作状态。当收到上位机发送的状态查询命令帧后,芯片会自动向上位机发送当前芯片状态的回传。上位机根据芯片状态的回传数据来判断当前芯片是处于空闲状态还是文本合成状态。

8.3 文本缓存功能支持的控制命令

芯片内部设有 16 段文本缓存区,段位分别为 0—15,用户可以把多条文本(≤ 16 条)发送到芯片的缓存区内,并在发送时指定存放的段位。存放后可以通过发送“缓存文本播放命令”来播放缓存的文本,并可以设置播报次数。

8.3.1 命令帧定义

名称	长度	说明
帧头	1 Byte	定义为十六进制“0xFD”
数据区长度	2 Bytes	用两个字节表示,高字节在前,低字节在后
数据区	小于 4k + 2 Bytes	命令字和命令参数,长度和“数据区长度”一致

表格 17 命令帧说明

命令帧中的数据区是由命令字和命令参数组成的,上位机使用命令字来实现文本缓存及播放的功能。

名称	发送的数据	说明
命令字	0x31	发送缓存文本命令
	0x32	播放缓存文本命令
命令参数	不同命令字有不同参数列表,详见各命令字说明	

表格 18 数据区命令字和命令参数

8.3.2 发送缓存文本命令

名称	发送的数据	说明
命令字	0x31	发送缓存文本命令
参数列表	1 字节的数据	设置本次文本应存储的起始缓存区段为 X
Data	缓存的文本	本次缓存的文本二进制内容 (设文本长度=Y) ($Y \leq (16-X) * 256$)
命令帧格式结构	帧头	数据区长度
	高字节	低字节
	0x00	0x01
	命令字	缓存区域
	0x31	0Xxx
	要缓存的文本

表格 19 发送缓存文本命令

注意:

- 缓存空间总共为 4K，共分 16 个区，每个区的空间为 256 字节。
- 设本次命令设置的起始缓存区段= X ($0 \leq X \leq 15$)，则本次发送的文本长度不能大于 $(16-X) * 256$ 字节。多余的文本将丢弃。
- 用户在发送【播放缓存文本命令 0x32】之前，可多次发送【发送缓存文本命令 0x31】任意安排区段的内容。但切记后来发送的文本不能部分覆盖或全部覆盖之前的文本，否则不保证播放的正确性。

8.3.3 播放缓存文本命令

名称	发送的数据	说明
命令字	0x32	播放缓存文本命令
参数列表	1 字节	高 4 位，取值 1 至 15，设置重复播放次数 低 4 位，设置文本编码格式 0x0 GB2312 0x1 GBK 0x2 BIG5 0x3 UNICODE
命令帧格式结构	帧头	数据区长度
	高字节	低字节
	0x00	0x01
	命令字	参数 1 字节
	0x32	高 4 位
		低 4 位
	0xX	0xX

表格 20 缓存文本播放命令

注意：发送缓存文本命令与播放缓存文本命令要按顺序执行，中间不要插入其它命令，若插入其

它命令，文本缓存区的内容可能被清空。

8.3.4 发送及播放缓存文本的举例

8.3.4.1 发送缓存文本命令

将文本“段缓存后一起播放。”存储在播放缓存的第3区段：

帧头	数据区长度		数据区		
	高字节	低字节	命令字	缓存区域	待缓存文本（段缓存后一起播放。）
0xFD	0x00	0x14	0x31	0x03	0xB6 0xCE 0xBB 0xBA 0xB4 0xE6 0xBA 0xF3 0xD2 0xBB 0xC6 0xF0 0xB2 0xA5 0xB7 0xC5 0xA1 0xA3

表格 21 文本缓存举例（1）

将文本“打乱顺序的一段文”存储在播放缓存的第1区段：

帧头	数据区长度		数据区		
	高字节	低字节	命令字	缓存区域	要缓存文本（打乱顺序的一段文）
0xFD	0x00	0x12	0x31	0x01	0xB4 0xF2 0xC2 0xD2 0xCB 0xB3 0xD0 0xF2 0xB5 0xC4 0xD2 0xBB 0xB6 0xCE 0xCE 0xC4

表格 22 文本缓存举例（2）

将文本“此功能可以将被”存储在播放缓存的第0区段：

帧头	数据区长度		数据区		
	高字节	低字节	命令字	缓存区域	待缓存文本（此功能可以将被）
0xFD	0x00	0x10	0x31	0x03	0xB4 0xCB 0xB9 0xA6 0xC4 0xDC 0xBF 0xC9 0xD2 0xD4 0xBD 0xAB 0xB1 0xBB

表格 23 文本缓存举例（3）

将文本“本按客户希望的顺序分”存储在播放缓存的第2区段：

帧头	数据区长度		数据区		
0xFD	高字节	低字节	命令字	缓存区域	待缓存文本（本按客户希望的顺序分）

	0x00	0x16	0x31	0x03	0xB1 0xBE 0xB0 0xB4 0xBF 0xCD 0xBB 0xA7 0xCF 0xA3 0xCD 0xFB 0xB5 0xC4 0xCB 0xB7 0xD0 0xF2 0xB7 0xD6
--	------	------	------	------	--

表格 24 文本缓存举例（4）

8.3.4.2 文本缓存播放命令

将编码格式为“GBK”的文本缓存里的内容“此功能可以将被打乱顺序的一段文本按客户希望的顺序分段缓存后一起播放。”重复播放 3 次。

帧头	数据区长度		数据区		
0xFD	高字节	低字节	命令字	参数 1 字节	
				高 4 位	低 4 位
	0x00	0x01	0x32	0x3	0x1

表格 25 文本缓存播放举例

8.4 语音编解码功能支持的控制命令

8.4.1 命令帧定义

名称	长度	说明
帧头	1 Byte	定义为十六进制“0xFD”
数据区长度	2 Bytes	用两个字节表示，高字节在前，低字节在后
数据区	小于 4k Bytes	命令字和命令参数，长度和“数据区长度”一致

表格 26 命令帧说明

上位机通过 UART 接口和 XFS5152CE 芯片连接,并且波特率设置为 115200bps 时,可以实现语音编解码功能。

数据区是由命令字和命令参数组成的，上位机使用命令字来实现语音合成芯片的各种功能。

数据区	命令字		命令参数
	命令字	说明	
	0x41	开始语音编码的命令	3 Bytes 的参数。参数 1：采样率；参数 2：设置比特率等级；参数 3：设置音量。
	0x42	开始语音解码的命令	3 Bytes 的参数。参数 1：采样率；参数 2：设置比特率等级；参数 3：设置音量。
	0x43	解码帧数据发送	上位机根据指定的压缩等级发送一帧数据给语音芯片进行解码

	0x44	停止录音或解码命令	无参数
--	------	-----------	-----

表格 27 数据区命令字和命令参数的说明

8.4.2 开始语音编码的命令

名称	发送的数据	说明					
命令字	0x41	开始编码命令					
参数列表 (3 个参数)	参数 1 (1 Byte)	设置采样率	设置为: 00, 采样率 8k				
			设置为: 01, 采样率 16k				
	参数 2 (1 Byte)	设置比特率等级	设置为: 00, 比特率等级 0 级				
			设置为: 01, 比特率等级 1 级				
			设置为: 02, 比特率等级 2 级				
			设置为: 03, 比特率等级 3 级				
			设置为: 04, 比特率等级 4 级				
			设置为: 05, 比特率等级 5 级				
	参数 3 (1 Byte)	设置音量	设置为: 00, 音量 0 级				
			设置为: 01, 音量 1 级				
			设置为: 02, 音量 2 级				
			设置为: 03, 音量 3 级				
			设置为: 04, 音量 4 级				
			设置为: 05, 音量 5 级				
			设置为: 06, 音量 6 级				
			设置为: 07, 音量 7 级				
			设置为: 08, 自动增益控制				
命令帧格式结构		帧头	数据区长度		数据区		
		0xFD	高字节 0x00	低字节 0x04	命令字 0x41	参数 1 0xXX	参数 2 0xXX
示例		如设置 16k 采样率、5 级比特率、音量采用自动增益控制, 发送的数据帧如下:					
		0xFD	0x00	0x04	0x41	0x01	0x05

表格 28 开始编码的命令说明

8.4.3 开始语音解码的命令

名称	发送的数据	说明
----	-------	----

命令字	0x42	开始解码命令				
参数列表 (3 个参数)	参数 1 (1 Byte)	设置采样率	设置为: 00, 采样率 8k			
			设置为: 01, 采样率 16k			
	参数 2 (1 Byte)	设置比特率等级	设置为: 00, 比特率等级 0 级			
			设置为: 01, 比特率等级 1 级			
			设置为: 02, 比特率等级 2 级			
			设置为: 03, 比特率等级 3 级			
			设置为: 04, 比特率等级 4 级			
			设置为: 05, 比特率等级 5 级			
	参数 3 (1 Byte)	设置音量	设置为: 00, 音量 0 级			
			设置为: 01, 音量 1 级			
			设置为: 02, 音量 2 级			
			设置为: 03, 音量 3 级			
			设置为: 04, 音量 4 级			
			设置为: 05, 音量 5 级			
			设置为: 06, 音量 6 级			
			设置为: 07, 音量 7 级			
命令帧格式结构	0xFD	帧头	数据区长度		数据区	
		高字节	低字节	命令字	参数 1	参数 2
		0x00	0x04	0x41	0xXX	0xXX
示例	如设置 16k 采样率、5 级比特率、5 级音量, 发送的数据帧如下:					
	0xFD	0x00	0x04	0x41	0x01	0x05

表格 29 开始解码的命令说明

下表以比特率等级为 0 (最低等级)、3(中间等级)、5 (最高等级) 三种情况为例, 说明在设置某个特定的录音采样率和比特率等级的情况下, 对应的每分钟音频采样数据量、编码压缩比和音频编码后的数据量。

采样率/采样精度	比特率等级	音频采样数据量 (PCM 格式,每分钟)	编码压缩比	音频编码后数据量 MB (每分钟)
8K/16Bit	0 (最低等级)	0.9MB	30	0.03
	3	0.9MB	13	0.07MB
	5 (最高等级)	0.9MB	10	0.09MB
16K/16Bit	0 (最低等级)	1.8MB	45	0.04MB
	3	1.8MB	22.5	0.08MB
	5 (最高等级)	1.8MB	14	0.13MB

表格 30 采样率比特率等级与压缩比和数据量对照表

例如: 选择设置采样率 8k、采样精度 16bit、比特率等级 05 时, 芯片每分钟的采样数据是

0.9MB，芯片采用 10 倍压缩比进行编码，编码后每分钟的数据量是 0.09MB，即这种情况下每分钟每分钟在芯片数据输出端需要进行传送的数据是 0.09MB。用户可据此选择合适的通讯速率或确定需要的存储空间。

8.4.4 发送数据到芯片进行解码

名称	发送的数据		说明			
命令字	0x43		根据指定的压缩等级发送一帧数据给芯片解码；与申请数据帧（0xFC 0x00 0x01 0x23）配合使用			
命令帧格式结构						
		帧头	数据区长度		数据区	
		0xFD	高字节	低字节	命令字	待解码数据
0xFF	0xFF		0x43		
示例		如：发送一帧采样率 8k、比特率等级 05 的数据，字节数对应为 62(16 进制为 3F)，发送的数据帧如下：				
		0xFD	0x00	0x3F	0x43

表格 31 发送解码数据的帧格式

8.4.5 停止语音编解码的命令

名称	发送的数据		说明				
命令字	0x44		停止语音编码或解码				
命令帧格式结构		0xFD	帧头	数据区长度		数据区	
			高字节	低字节	命令字		
			0x00	0x01	0x44		

表格 32 停止编解码的命令说明

8.4.6 编解码相关的状态回传

回传数据类型	回传数据					触发条件
返回编码后数据给主机	帧头	数据区长度		数据区		上位机发送启动编码命令
	0xFC	高字节 (0Xxx)	低字节 (0xXX)	命令字 (0x20)	编码数据	

录音错误	0xFC	0X00	0X01	0X21	
录音完毕	0xFC	0X00	0X01	0X22	
申请数据	0xFC	0X00	0X01	0X23	向主机申请待解码数据(注 1)

表格 33 编解码相关的状态回传

注意：解码数据需按帧发送，主机接收到申请数据的请求后发送一帧数据。

8.4.7 编解码延迟时间参数

采样率	缓冲时间(ms)	编码延时(ms)	解码延时(ms)
8K	80	16	10
16K	40	16	10

表格 34 编解码延迟时间参数

8.5 语音识别功能支持的控制命令

8.5.1 命令帧定义

名称	长度	说明
帧头	1 Byte	定义为十六进制“0xFD”
数据区长度	2 Bytes	用两个字节表示，高字节在前，低字节在后
数据区	小于 4k Bytes	命令字和命令参数，长度和“数据区长度”一致

表格 35 命令帧说明

命令帧的数据区是由命令字和命令参数组成的，上位机使用命令字来实现语音识别功能。

名称	发送的数据	说明
命令字	0x10	启动语音识别命令
	0x1F	停止语音识别命令，没有参数
命令参数	不同命令字有不同参数列表，详见各命令字说明	

表格 36 数据区命令字和命令参数

8.5.1 启动语音识别命令

名称	发送的数据	说明
----	-------	----

命令字	0x10	启动语音识别功能	
参数列表	无		
命令帧格式结构	帧头	数据区长度	数据区
	0xFD	高字节	低字节
		0x00	0x01

表格 37 启动语音识别命令

8.5.2 停止语音识别命令

名称	发送的数据	说明
命令字	0x1F	停止当前的语音识别
参数列表	无	
命令帧格式结构	帧头	数据区长度
	0xFD	高字节
		0x00

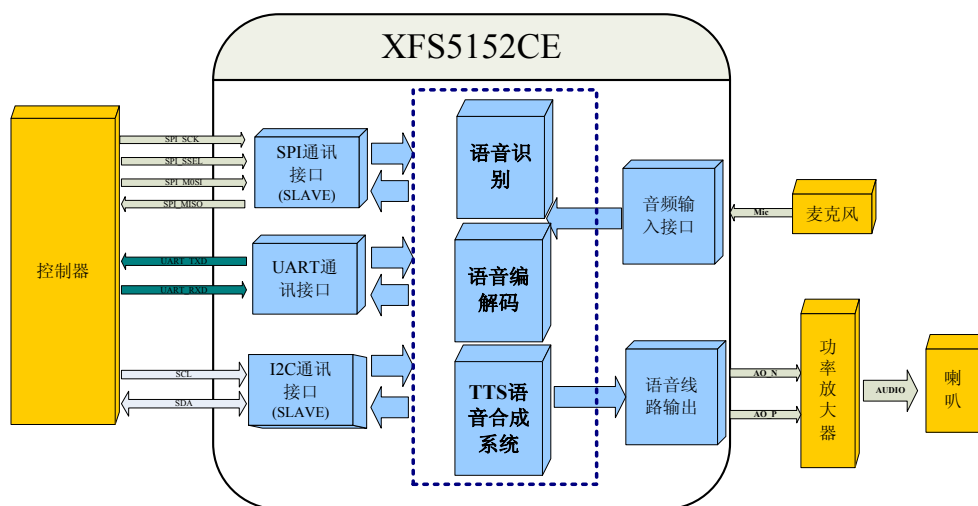
表格 38 停止语音识别命令

8.5.3 芯片语音识别结果回传

回传数据类型	回传数据					触发条件					
返回识别结果 ID	帧头	数据区长度		数据区		上位机启动识别命令后，有识别结果产生时，最长检测到语音 4 秒会出结果或者产生拒识别					
	0xFC	高字节 (0Xxx)	低字节 (0xXX)	命令字 (0x01)	识别结果 (0xXX)						
语音识别系统出现超时	<table><tr><td>0xFC</td><td>0X00</td><td>0X01</td><td colspan="2">0X02</td></tr></table>					0xFC	0X00	0X01	0X02		上位机启动识别命令后，检测无人说话时，5 秒
0xFC	0X00	0X01	0X02								
语音识别系统出现拒识	<table><tr><td>0xFC</td><td>0X00</td><td>0X01</td><td colspan="2">0X03</td></tr></table>					0xFC	0X00	0X01	0X03		上位机启动识别命令后，无结果时
0xFC	0X00	0X01	0X03								
语音识别系统出现内部错误	<table><tr><td>0xFC</td><td>0X00</td><td>0X01</td><td colspan="2">0X04</td></tr></table>					0xFC	0X00	0X01	0X04		上位机启动识别命令后，资源出错时
0xFC	0X00	0X01	0X04								

表格 39 芯片识别结果回传

9 语音芯片系统结构图



图表 7 XFS5152CE 语音合成系统构成框图

一般应用中语音合成系统最小系统需要包括：控制器模块、XFS5152CE 芯片、功放模块、喇叭。如果需要使用语音识别功能、或者语音编解码功能，系统中还需要增加麦克风。

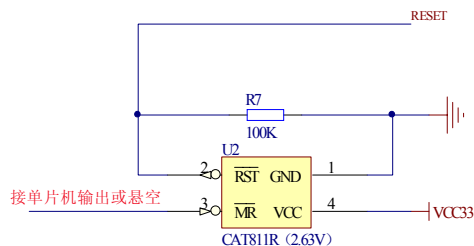
语音合成系统中，主控制器和 XFS5152CE 芯片之间可以通过 UART 接口、或者 I2C 接口、或者 SPI 接口连接，控制器可通过上述通讯接口向 XFS5152CE 芯片发送控制命令和文本，XFS5152CE 芯片接收到文本后合成为语音信号输出，输出的信号经功率放大器进行放大后连接到喇叭进行播放。

用户在使用语音识别功能时，上位机发送启动语音识别功能的命令给语音芯片，芯片把从麦克风采集到的语音数据，通过内部的识别模块进行转换成相应的识别结果，通过通讯接口回传给控制器。

用户在使用语音编解码功能时（通讯接口必须选择 UART 接口，并且波特率设置为 115200bps），上位机发送启动编解码的命令给语音芯片，芯片内部的语音编解码模块把采集到的音频数据进行编码并通过 UART 接口实时传送给上位机，或者对上位机传送来的音频数据进行解码并实时播放出来。

图表 10 晶振参考电路

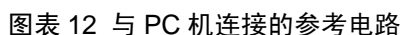
XFS5152CE 芯片的复位条件是： \overline{RESET} 引脚置低电平，时间持续 40ms 以上。可以通过 RC 阻容复位电路 (R=100K, C=2.2uF)，或者 CAT811 复位芯片对 XFS5152CE 进行复位。下图是 CAT811 复位模块的电路图。



图表 11 复位参考电路

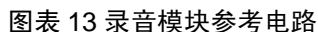
下面对 PC 机、单片机两种上位机和芯片的连接方法分别进行说明:

XFS5152CE 芯片与 PC 通过 UART 串口相连，需要使用 MAX3221 电平转换模块进行电平的转换，如下图所示。



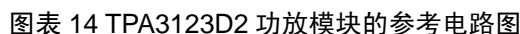
可通过 UART 接口、I2C 接口、SPI 接口相连。连接方法见图表 1、图表 3 和图表 5。

在应用中如果需要使用到录音功能，麦克风模块可以参考以下电路进行设计，MIC 偏置电压 MIC_BIAS 由芯片 12 脚输出。下图电路中网络标号 MIC 为麦克风接入处，MIC_N、MIC_P 连接到芯片的 13、15 脚，网络标号为 MIC VCC 的两处互相相连即可。

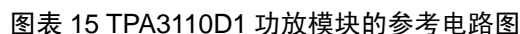


本文中提供三种音频功放芯片 TPA3123D2、TPA3110D1、TPA2005D1 的连接方法供参考,关于功放芯片的具体特性请参考其官方提供的数据手册。

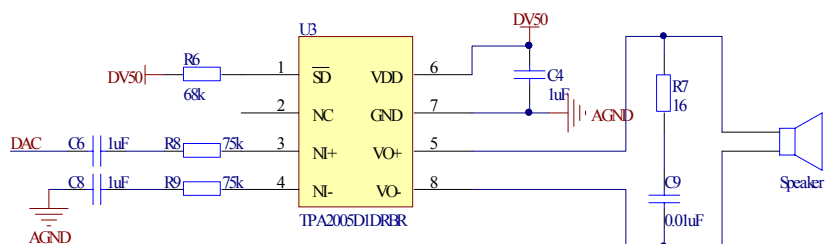
TPA3123D2 的工作电压范围 10V~30V,最大输出功率为 25W。下图为 TPA3123D2 参考电路。



TPA3110D1 的工作电压范围 8V~26V，最大输出功率为 15W。下图为 TPA3110D1 的参考电路。



VCC 电压范围 2.5~5.5V，最大输出功率 1.4W。下图为 TPA2005D1 的参考电路。



图表 16 TPA2005D1 功放模块的参考电路图

11 示例程序

11.1 发送语音合成文本的示例程序

下面以 51 单片机作为上位机为例，用 C51 语言实现一段文本合成的程序示例，假设要合成的文本内容为：“欢迎使用科大讯飞语音合成芯片。”

```
#include <reg51.h>
#include <string.h>

void main(void)
{
    /*****需要发送的文本*****/
    char code szText[] = {"欢迎使用科大讯飞语音合成芯片。"};
    unsigned int nLength = 0;
    int i = 0;

    /*****串口的初始化*****/
    TL1    = 0xFA;    // 在 11.0592MHz 下，设置波特率 9600bps，工作方式 2
    TH1    = 0xFA;
    TMOD   = 0x20;
    SCON   = 0x50;    // 串口工作方式 1，允许接收
    PCON   = 0x80;
    EA     = 0;
    REN    = 1;
    TI     = 0;       // 发送中断标志位置零
    RI     = 0;       // 接收中断标志位置零
    TR1    = 1;       // 定时器 1 用作波特率发生

    /*****发送过程*****/
    nLength = strlen(szText); // 需要发送文本的长度
    SBUF = 0xFD;              // 向串口发送帧头标志
    while ( TI == 0 );        // 等待发送中断标志位置位
    TI = 0;                   // 发送中断标志位清零
    SBUF = 0x00;              // 发送待合成文本长度的高字节
    while ( TI == 0 );
    TI = 0;
    SBUF = nLength + 2;        // 发送待合成文本长度的低字节
    while ( TI == 0 );
    TI = 0;
```

```
SBUF = 0x01;           // 发送文本合成命令字
while ( TI == 0 );
TI = 0;
SBUF = 0x00;           // 发送文本编码格式
while ( TI == 0 );
TI = 0;
for ( i = 0; i < nLength; i++ ) // 依次发送待合成的文本数据
{
    SBUF = szText[i];
    while ( TI == 0 );
    TI = 0;
}
while ( 1 );           // 死循环，阻止程序退出
}
```

以上为发送一帧 TTS 文本数据的程序模块。发送后可接收到语音芯片自动发送的回传，如果先后收到“41”和“4F”，则说明文本被正确接收，且表明合成完毕，芯片处于空闲状态；如果收到的是“45”，则说明数据帧错误，需要重新发送或复位。

注意：上面的 Demo 主要说明了发送过程需要遵循的协议问题，实际系统中还需要有波特率的设置程序；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。

11.2 语音编解码的示例程序

```
#include "StdAfx.h"

#include "Codec.h"

static FILE *Fp;

static CODECMODEID g_systate;

static VOLUMEID g_volume;

//存储数据回调
//录音编码时候存储压缩数据

bool app_savedata(
    void*          pSrcBuffer,
```

```
int      nBufferBytes
)
{
    //windows 平台存储文件
    fwrite(pSrcBuffer, 1, nBufferBytes, Fp);

    return 0;
}
//读取数据回调
//解码播音时候读取压缩数据
bool app_loadsave(
    void*      pDstBuffer,
    int*       pnBufferBytes
)
{
    //windows 平台读取文件
    *pnBufferBytes = fwrite(pDstBuffer, 1, *pnBufferBytes, Fp);
    return 0;
}
//送命令回调
//负责将编解码的命令通过串口发送出去
bool app_sendcommand(
    void*      pSrcBuffer,
    int        nBufferBytes
)
{
    //根据具体的平台，将 buf 中数据用串口发送出去

    return 0;
}
```

```
}  
  
//获取串口接收数据  
  
char UART_Get_byte(void)  
{  
  
    char ch;  
  
    //从串口接收 buf 中取一个数据  
  
    return ch;  
}  
  
// 启动录音编码  
  
void StartEncode(void)  
{  
  
    VOLUMEID volume = BR_volume7;  
  
    g_systate = Codec_encode;  
  
    g_volume = BR_volume7;  
  
    Fp = fopen("codec.bts", "wb");  
  
    //开始编码,设置采样率 16k, 比特率 5, 音量 7  
  
    Codec_Start(g_systate, SR_16000, BR_bitrate5, g_volume, NULL, app_savedata,  
app_sendcommand);  
  
    while (1)  
    {  
  
        //如果数据处理字节返回非 0, 有错误, 退出  
  
        if (Codec_Byte_Proc(UART_Get_byte()))  
        {  
  
            Codec_Stop();  
  
        }  
  
        //是否需要退出  
  
        if (g_systate != Codec_encode)  
        {
```



```
        Codec_Stop();

    }

    //是否重设音量

    if (g_volume != volume)

    {

        g_volume = volume;

        Codec_SetVolume(g_volume);

    }

}

fclose(Fp);
}

void StartDecode(void)
{

    VOLUMEID volume = BR_volume7;

    // 解码播音样例

    g_systate = Codec_decode;

    g_volume = BR_volume7;

    Fp = fopen("codec.bts", "rb");

    Codec_Start(g_systate, SR_16000, BR_bitrate5, g_volume, app_loadsave, NULL,
app_sendcommand);

    while (1)

    {

        //如果数据处理字节返回非 0，有错误或播音完成，退出

        if (Codec_Byte_Proc(UART_Get_byte()))

        {

            Codec_Stop();

        }

    }

}
```

```
//是否需要退出
if (g_systate != Codec_encode)
{
    Codec_Stop();
}
//是否重设音量
if (g_volume != volume)
{
    g_volume = volume;
    Codec_SetVolume(g_volume);
}
}
fclose(Fp);
}
```

注意：编码的内部函数 `codec.c` 没有包含在上面的参考程序中实现，已作为参考程序另外提供给用户。

12 文本控制标记

12.1 文本控制标记的使用方法

XFS5152CE 芯片的语音合成功能支持多种文本控制标记,可以满足用户对语音合成发音人、音量、语速、语调等的设置。

文本控制标记的格式一般是半角中括号（即“[]”）内一个小写字母、一个阿拉伯数字，如：[m3]，标记的使用方法和合成文本完全一致。通讯协议详见本开发指南的“8.2.1 语音合成命令”章节。

用户可以把标记作为文本单独发送到芯片上，如：只发送 “[v3]” 到芯片上设置合成音量为 3 级，或者把标记和其他要合成的文本放在一起发送给芯片上,如：“[v3]我在小声说话, [v10]我在大声说话”。

标记只是作为控制标记实现设置功能，不会合成为声音输出。如：“[s1]我慢条斯理。[s8]我快言快语”中，经过标记的设置，前一句合成语速会很慢，后一句合成语速会很快，但不会读出“s1”和“s8”。

12.2 文本控制标记列表

作用	控制	详细说明		芯片
合成风格设置	[f?]	? 为 0，一字一顿的风格		[f1]
		? 为 1，正常合成		
合成语种设置	[g?]	? 为 0，自动判断语种		[g1]
		? 为 1，阿拉伯数字、度量单位、特殊符号等合成为中文		
		? 为 2，阿拉伯数字、度量单位、特殊符号等合成为英文		
设置单词的发音方式	[h?]	? 为 0，自动判断单词发音方式		[h1]
		? 为 1，字母发音方式		
		? 为 2，单词发音方式		
设置对汉语拼音的识别	[i?]	? 为 0，不识别汉语拼音		[i0]
		? 为 1，将“拼音+1 位数字（声调）”识别为汉语拼音，例如：hao3		
选择发音人	[m?]	中英文发音人	? 为 3，设置发音人为小燕(女声，推荐发音人)	[m3]
			? 为 51，设置发音人为许久(男声，推荐发音人)	
			? 为 52，设置发音人为许多(男声)	
			? 为 53，设置发音人为小萍(女声)	

		? 为 54, 设置发音人为唐老鸭(效果器) ? 为 55, 设置发音人为许小宝(女童声)	
设置数字处理策略	[n?]	? 为 0, 自动判断 ? 为 1, 数字作号码处理 ? 为 2, 数字作数值处理	[n0]
数字“0”在读作英文、号码时的读法	[o?]	? 为 0, 读成“zero” ? 为 1, 读成“欧”音	[o0]
合成过程中停顿一段时间	[p?]	? 为无符号整数, 表示停顿的时间长度, 单位为毫秒(ms)	
设置姓名读音策略	[r?]	? 为 0, 自动判断姓氏读音 ? 为 1, 强制使用姓氏读音规则	[r0]
设置语速	[s?]	? 为语速值, 取值: 0~10	[s5]
设置语调	[t?]	? 为语调值, 取值: 0~10	[t5]
设置音量	[v?]	? 为音量值, 取值: 0~10	[v5]
设置提示音处理策略	[x?]	? 为 0, 不使用提示音 ? 为 1, 使用提示音	[x1]
设置号码中“1”的读法	[y?]	? 为 0, 合成号码“1”时读成“幺” ? 为 1, 合成号码“1”时读成“一”	[y0]
是否使用韵律标记“*”和“#”	[z?]	? 为 0, “*”和“#”读出符号 ? 为 1, 处理成韵律, “*”用于断词, “#”用于停顿	[z0]
为单个汉字强制指定拼音	[=?]	? 为标记前一个汉字的拼音+声调(1~5 分别表示阴平, 阳平, 上声, 去声和轻声) 5 个声调。例如: “着[=zhuo2]手”, “着”字读作“zhuó”	
恢复默认的合成参数	[d]	所有设置(除发音人设置、语种设置外)恢复为默认值	

表格 40 控制标记列表

注意:

- 所有的控制标识均为半角字符。
- 控制标识需要按照语音合成命令的格式发送, 控制标记作为文本进行合成, 即合成命令是“帧头 + 数据区长度 + 合成命令字 + 文本编码格式 + 控制标记文本”的格式。
- 控制标识为全局控制标识, 也就是只要用了一次, 在不对芯片进行复位、或断电、或使用[d]恢复默认设置的条件下, 其后发送给芯片的所有文本都会处于它的控制之下。
- 当芯片掉电或是复位后, 原来的设置过的标识会失去作用, 芯片将恢复到所有的默认值。

12.3 控制标记使用示例

12.3.1 [f?]标记的使用

示例文本	芯片解释
[f0]科大讯飞	合成为：“科-大-讯-飞”
[f1]科大讯飞	合成为：“科大讯飞”

表格 41 [f?]标记的使用

12.3.2 [g?] 标记的使用

示例文本	芯片解释
[g0]123km	合成为：“一百二十三公里”
[g1]123km	合成为：“一百二十三公里”
[h0][g2] 123km	合成为：“one hundred and twenty three kilometres”

表格 42 [g?]标记的使用

12.3.3 [h?] 标记的使用

示例文本	芯片解释
[h0] 你的英语是 “you”	合成为：“你的英语是 you”
[h1]你的英语是 “you”	合成为：“你的英语是 y-o-u”
[h2] 你的英语是 “you”	合成为：“你的英语是 you”

表格 43 [h?]标记的使用

12.3.4 [i?]标记的使用

示例文本	芯片解释
[i0]科大讯 fei1	芯片不识别拼音格式。合成为：“科大讯 F E I”
[i1]科大讯 fei1	芯片识别拼音格式。合成为：“科大讯飞”

表格 44 [i?]标记的使用

注意：[i?]标记设置为[i1]时，会对英语单词、字母缩写、提示音等的合成造成影响，建议使用后

及时使用[d]标记恢复成默认设置。

12.3.5 [m?]标记的使用

示例文本	芯片解释
[m3]我是晓燕。	女声合成：“我是晓燕”
[m51]我是许久。	男声合成：“我是许久”
[m52]我是许多。	男声合成：“我是许多”
[m53]我是晓萍。	女声合成：“我是晓萍”
[m54]我是唐老鸭。	唐老鸭模拟声合成：“我是唐老鸭”
[m55]我是许小宝。	女童声合成：“我是许小宝”

表格 45 [m?]标记的使用

12.3.6 [n?]标记的使用

示例文本	芯片解释
[n0]234343545	芯片自动判断。读作：二三四三四五四五
[n1]234343545	芯片强制按照号码的方式合成数字串。读作：二三四三四三五四五
[n2]234343545	芯片强制按照数值的方式合成数字串。读作：两亿三千四百三十四万三千五百四十五

表格 46 [n?]标记的使用

12.3.7 [o?]标记的使用

示例文本	芯片解释
[g2][h0][n1][o0]8016700	读作：eight o(音“欧”) one, six seven o o(音“欧”)
[g2][h0][n1][o1]8016700	读作：eight zero one, six seven zero zero

表格 47 [o?]标记的使用

12.3.8 [p?]标记的使用

示例文本	芯片解释
欢迎使用[p500]科大讯飞[p1000]语音合成芯片	芯片在“欢迎使用”合成完毕后静音 500 毫秒，在“讯飞”合成完毕之后静音 1000 毫秒。“p”后面所带的整数越大，静音的时间越长，最大的静音长度为 268 秒。

表格 48 [p?]标记的使用

12.3.9 [r?]标记的使用

示例文本	芯片解释
[r0] 查丽	芯片按照默认的方式合成。读作：chá lì
[r1] 查丽	芯片强制的将文本的第一个汉字按照姓氏的读音合成。读作：zhā lì

表格 49 [r?]标记的使用

12.3.10 [s?]标记的使用

示例文本	芯片解释
您好, [s9]欢迎使用科大讯飞[s5]语音合成芯片	芯片语速的调节功能, 提供 11 级的语速调节, 芯片默认的语速为 5。本句合成时“您好”为 5 级语速, “欢迎使用科大讯飞”为 9 级语速, “语音合成芯片”为 5 级语速。

表格 50 [s?]标记的使用

12.3.11 [t?]标记的使用

示例文本	芯片解释
您好, [t10]欢迎使用科大讯飞[t0]语音合成芯片	芯片语调的调节功能, 提供 11 级的语调调节, 芯片默认的语调大小为 5, 最小为 0, 最大为 10。本句合成时“您好”为 5 级语调, “欢迎使用科大讯飞”为 10 级语调, “语音合成芯片”为 0 级语调。

表格 51 [t?]标记的使用

12.3.12 [v?]标记的使用

示例文本	芯片解释
欢迎[v3]使用[v8]科大讯飞语音合成芯片	芯片音量的调节功能, 提供 11 级的音量调节, 最小 0 为静音, 芯片默认的音量大小为 5。本句合成时“欢迎”为 5 级音量, “使用”为 3 级音量, “科大讯飞语音合成芯片”为 8 级音量。

表格 52 [v?]标记的使用

12.3.13 [x?]标记的使用

示例文本	芯片解释
提示音铃声[x0]sound101 对应的声音是[x1]sound101	[x0]和[x1]均起作用。读作：预先制定的提示音铃声 sound101 为（提示音 sound101 的铃声）

表格 53 [x?]标记的使用

12.3.14 [y?]标记的使用

示例文本	芯片解释
[y0]010-58301005-8016	芯片按照“幺”的读法合成号码文本中的“1”。读作：零幺零 五八三零 幺零零五 转 八零幺六
[y1]010-58301005-8016	芯片按照“一”的读法合成号码文本中的“1”。读作：零一零 五八三零 一零零五 转 八零一六

表格 54 [y?]标记的使用

注意：此标记必须是在合成号码类型文本的时候才有效。

12.3.15 [z?]标记的使用

示例文本	芯片解释
[z0]安徽#科大*讯飞	芯片读成“安徽井号科大星号讯飞”
[z1]安徽#科大*讯飞	芯片读成“安徽 科大 讯飞”，#和*的地方被处理成短暂的停顿

表格 55 [z?]标记的使用

12.3.16 [=?]标记的使用

示例文本	芯片解释
曾[=deng1]国藩	芯片强制的将“曾”按照“dēng”音合成播报

表格 56 [=?]标记的使用

12.3.17 [d]标记的使用

示例文本	芯片解释
[s2][v3][t10]科大讯飞, [d]科大讯飞	前一个“科大讯飞”按照 2 级语速、3 级音量、10 级语调合成；

	后一个“科大讯飞”恢复到默认的语速、音量、语调进行合成。
--	------------------------------

表格 57 [d]标记的使用

13 提示音

13.1 芯片集成提示音

芯片集成了 80 首声音提示音，可以依据使用场合选用作为信息提示音。

13.1.1 信息提示音列表

信息提示音(共 25 首)				
sound101	sound102	sound103	sound104	sound105
sound106	sound107	sound108	sound109	sound110
sound111	sound112	sound113	sound114	sound115
sound116	sound117	sound118	sound119	sound120
sound121	sound122	sound123	sound124	sound125

表格 58 信息提示音列表

13.1.2 铃声提示音列表

铃声提示音(共 25 首)				
sound201	sound202	sound203	sound204	sound205
sound206	sound207	sound208	sound209	Sound210
sound211	sound212	Sound213	sound214	sound215
sound216	sound217	sound218	sound219	Sound220
Sound221	sound222	sound223	sound224	sound225

表格 59 铃声提示音列表

13.1.3 警报提示音列表

警报提示音(共 30 首)				
sound301	sound303	sound303	sound304	sound305
sound306	sound307	sound308	sound309	sound310
sound311	sound312	sound313	sound314	sound315
sound316	sound317	sound318	sound319	sound320
sound321	sound322	sound323	sound324	sound325

sound326	sound327	sound328	sound329	sound330
----------	----------	----------	----------	----------

表格 60 警报提示音列表

注意：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或数字时，需要使用标点符号、空格、回车等与其他字母隔开，芯片才能够识别提示音。

13.2 提示音定制功能

XFS5152CE 芯片预留了 200K 空间，对批量采购的客户可提供提示音添加、定制的服务。

用户提供音频文件的要求：

1、用户可以添加采样率采样精度分别为 16K16bit，wav 格式的音频文件，不是此格式的文件需要进行转换。

2、用户可以添加 16K16bit、wav 格式的音频文件的总量可达 4M（由于提示音在存放到芯片时会进行格式转换，压缩比例是 20：1，200K x 20 即为 4M）。

如果用户需要添加总量超过 4M 的提示音，可支持把内部提示音删除一部分，获得更多的存储空间。

注意：如有添加提示音的需求，需要提交与商务人员沟通，提供音频和填写相应的资料,并且增加提示音的名称不能与已有的提示音重合。

14 语音识别命令词

14.1 命令词列表

命令词(共 30 个)				
我在吃饭	我在修车	我在加油	正在休息	同意
不同意	我去	现在几点	今天几号	读信息
开始读	这是哪儿	打开广播	关掉广播	打开音乐
关掉音乐	再听一次	再读一遍	大声点	小声点
读短信	读预警信息	明天天气怎么样	紧急预警信息	开始
停止	暂停	继续读	确定开始	取消

表格 61 报警提示音列表

14.2 命令词定制说明

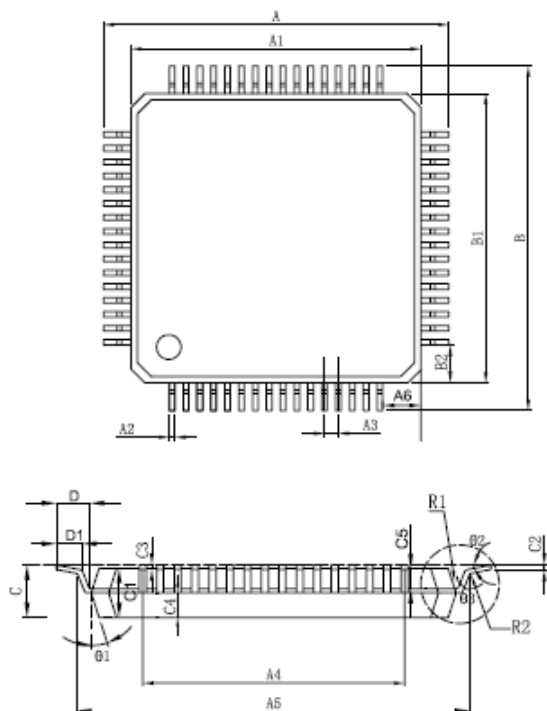
XFS5152CE 芯片可支持 30 个命令词的识别。芯片出默认设置的是 30 个车载、预警等行业常用识别命令词。客户如需要更改成其他的识别命令词，可进行命令词定制。

注意：如需要定制芯片的命令词，请在订货前和语音芯片的商务人员联系确认。确认满足定制条件的芯片用户，需要提交填写好的命令词定制需求单。为了保证较好的识别率，请尽量选择互相之间区分度比较大的识别命令词。

15 产品规格

15.1 封装

XFS5152CE 芯片封装形式为 LQFP64。具体信息如下：



尺寸 标注	最小 (mm)	最大 (mm)	尺寸 标注	最小 (mm)	最大 (mm)
A	11.80	12.20	C2	0.09	0.18
A1	9.9	10.10	C3	0.05	0.15
A2	0.17	0.25	C4	0.6365TYP	
A3	0.5TYP		C5	0.6365TYP	
A4	7.5TYP		D	0.9	1.1
A5	11.0TYP		D1	0.45	0.6
A6	1.15TYP		R1	0.15	
B	11.80	12.20	R2	0.15	
B1	9.90	10.10	Θ1	12° TYP	
B2	1.15TYP		Θ2	0° TYP	7° TYP
C	1.30	1.60	Θ3	4° TYP	
C1	1.35	1.45			

图表 17 芯片封装图

15.2 特性参数

15.2.1 极限值

参数	符号	最小值	最大值	单位
工作电压	PWR_IN	3.1	4.5	V
引脚输入电压	V _{IN}	-0.3	3.6	V
存储温度	T _{STO}	-55	125	°C
工作温度	T _A	-40	85	°C

表格 62 芯片电气极限参数

注意：如果超出表中所列的极限参数，将可能导致器件损坏。

15.2.2 直流电气特性

参数	符号	条件	最小值	典型值	最大值	单位
工作电压	PWR_IN		3.1		4.5	V
输入低电平	V _{IL}				0.3VDDIO	V
输入高电平	V _{IH}		0.7VDDIO			V
输出低电平	V _{OL}	取决于负载			0.4	V
输出高电平	V _{OH}	取决于负载	2.4			V
工作电流	I _{op1}	合成文本	29		32	mA
	I _{op1}	编码		38.2		
	I _{op1}	解码		32.3		
输入漏电流	I _I	V _I = 3.3V 或 0V	-1		1	uA
输出电流	I _{OH}			8.2		mA
吸收电流	I _{OL}			13.2		mA
复位时间	T _{RST}			40		ms
初始化时间	T _{INI}			240		ms

表格 63 电气直流特性参数

15.2.3 音频 DAC 特性

负载电阻 Ω	最大峰峰值电压 V	输出功率 mW
---------------	-----------	---------

600	1.1	1.07
10K	1.1	0.064

表格 64 音频 DAC 特性参数

15.2.4 非文本合成状态下的功耗参数

合成后空闲	编码后空闲	解码后空闲	省电
34mA	35.5mA	35.5 mA	5mA

表格 65 非文本合成状态下的功耗参数对应表

15.3 接收合成命令到开始播音间隔时间

接收文本的大小	典型值
4096 字节文本	530ms

表格 66 接收合成命令到开始播音间隔时间

15.4 焊接工艺要求

15.4.1 烘烤温度及时间

XFS5152CE 的潮敏等级是 3 级，在 $\leq 30^{\circ}\text{C}/60\%\text{RH}$ 的环境条件下，落地寿命为 168 小时。

产品名称	潮敏等级	落地寿命 (拆开包装后，在 $\leq 30^{\circ}\text{C}/60\%\text{RH}$ 的湿度条件下)
XFS5152CE	3	168 H

表格 67 潮敏分类等级及落地寿命

当拆装芯片时间超过落地寿命时，使用时需要按照下表的规定烘烤后，再进入回流焊流程。烘烤时间的规定如下：

芯片超过落地寿命 时间	>72 H	≤ 72 H	>72 H	≤ 72 H	>72 H	≤ 72 H
烘烤温度	125 $^{\circ}\text{C}$		90 $^{\circ}\text{C}$ $\leq 5\%\text{RH}$		40 $^{\circ}\text{C}$ $\leq 5\%\text{RH}$	
烘烤时间要求	9 H	7 H	33 H	23 H	312H	216H

表格 68 芯片烘烤的参考条件

注意：

- 1) 上表中：RH 表示相对湿度；H 表示小时；
- 2) 烘烤时料盘的耐温性能应符合要求。

15.4.2 回流焊的峰值温度

最低峰值温度	最高峰值温度
240℃	260℃

表格 69 回流焊的峰值温度

注意：根据焊剂融化点，可能要求采用更高的温度，锡膏的典型温度值：铅锡膏剂为 $220 \pm 5^{\circ}\text{C}$ ；锡银铜剂膏为 $245 \pm 5^{\circ}\text{C}$ ，具体需依据生产厂商的规格说明。

16 附录

16.1 GB2312 编码简介

GB 2312 码是中华人民共和国国家标准汉字信息交换用编码，全称《信息交换用汉字编码字符集 基本集》，标准号为 GB 2312—80（GB 是“国标”二字的汉语拼音缩写），由中华人民共和国国家标准总局发布，1981 年 5 月 1 日实施。习惯上称国标码、GB 码，或区位码。它是一个简化字汉字的编码，通行于中国大陆地区。新加坡等地也使用这一编码。

GB 2312-80 收录简化汉字及一般符号、序号、数字、拉丁字母、日文假名、希腊字母、俄文字母、汉语拼音符号、汉语注音字母，共 7445 个图形字符。其中汉字以外的图形字符 682 个，汉字 6763 个。

GB 2312-80 规定，“对任意一个图形字符都采用两个字节（Byte）表示。每个字节均采用 GB 1988—80 及 GB 2311—80 中的七位编码表示。两个字节中前面的字节为第一字节，后面的字节为第二字节。”习惯上称第一字节为“高字节”，第二字节为“低字节”。

16.2 UNICODE 编码简介

在创造 UNICODE 之前，有数百种编码系统。但是，没有任何一个编码可以包含足够的字符。例如，仅欧州共同体就需要好几种不同的编码来包括所有的语言。即使是单一的一种语言，如英语，也没有哪一个编码可以适用于所有的字母，标点符号，和常用的技术符号。

这些编码系统也会互相冲突。也就是说，两种编码可能使用相同的数字代表两个不同的字符，或使用不同的数字代表相同的字符。任何一台特定的计算机(特别是服务器)都需要支持许多不同的编码，但是，不论什么时候数据通过不同的编码或平台之间，那些数据总会有损坏的危险。

而在 UNICODE 标准中，提供了 1,114,112 个码点，不仅可以包含当今世界使用的所有语言文字和其他符号，也足够容纳绝大多数具有历史意义的古文字和符号。并且，UNICODE 给每个字符提供了一个唯一的数字，不论是什么平台，不论是什么程序，不论什么语言。

UNICODE 标准已经被工业界的领导们所采用，例如：Apple, HP, IBM, JustSystem, Microsoft, Oracle, SAP, Sun, Sybase, Unisys 等等。最新的标准都需要 UNICODE，例如 XML, Java, ECMAScript, LDAP, CORBA 3.0, WML 等等，并且 UNICODE 是实现 ISO/IEC 10646 的正规方式。许多操作系统，所有最新的浏览器和许多其他产品都支持它。UNICODE 标准的出现和支持它的工具的存在，是近来全球软件技术最重要的发展趋势。

16.3 GBK 编码简介

GB2312-80 仅收汉字 6763 个，这大大少于现有汉字，随着时间推移及汉字文化的不断延伸推广，有些原来很少用的字，现在变成了常用字，例如：朱镕基的“镕”字，未收入 GB2312-80，现在大陆的报业出刊只得使用（金+容）、（金容）、（左金右容）等来表示，形式不一而同，这使得表示、存储、输入、处理都非常不方便，对于搜索引擎等软件的构造来说也不是好消息，而且这种表示没有统一标准。

为了解决这些问题，以及配合 UNICODE 的实施，全国信息技术化技术委员会于 1995 年 12 月 1 日《汉字内码扩展规范》。GBK 向下与 GB2312 完全兼容，向上支持 ISO-10646 国际标准，在前者向后者过渡过程中起到了承上启下的作用。

GBK 是 GB2312-80 的扩展，是向上兼容的。它包含了 20902 个汉字，其编码范围是 0x8140-0xfefe，剔除高位 0x80 的字位。其所有字符都可以一对一映射到 UNICODE 2.0。

GBK 亦采用双字节表示，总体编码范围为 8140-FEFE 之间，首字节在 81-FE 之间，尾字节在 40-FE 之间，剔除 XX7F 一条线。

微软公司自 Windows 95 简体中文版开始支持 GBK 代码，标准叫法是 Windows codepage 936，也叫做 GBK（国标扩展），它也是 8-bit 的变长编码。

16.4 BIG5 编码简介

BIG5 是台湾计算机界实行的汉字编码字符集。它包含了 420 个图形符号和 13070 个繁体汉字（不包含简化汉字）。编码范围是 0x8140-0xFE7E、0x81A1-0xFEFE，其中 0xA140-0xA17E、0xA1A1-0xA1FE 是图形符号区，0xA440-0xF97E、0xA4A1-0xF9FE 是汉字区。