



XF-S4240 中文语音合成模块 开发指南

V 1.1 - 2008 年 9 月 10 日

安徽科大讯飞信息科技股份有限公司

安徽省合肥市国家级高新技术产业开发区信息产业基地讯飞语音大厦



86-0551-5331878/5397818



86-0551- 5331837



www.iflytek.com

版本历史

版 本	日 期	修改记录	作 者
Rev0.1	2005-12-7	DRAFT	科大讯飞
Rev0.2	2006-1-19	缺电路参考和文本使用范例	科大讯飞
Rev0.3	2006-1-20	添加了相关的参考电路	科大讯飞
Rev0.4	2006-1-21	添加发送示例修改电路图	科大讯飞
Rev0.5	2006-2-6	增加提示音指南及文档使用范例	科大讯飞
Rev0.6	2006-2-7	修正已知文字错误	科大讯飞
Rev0.7	2006-2-28	修改章节 1.2	科大讯飞
Rev0.8	2006-3-31	删除波特率中容易引起误解的部分	科大讯飞
Rev0.9	2006-6-6	添加 I2C 设备地址的描述	科大讯飞
Rev1.0	2007-12-12	修改应用参考电路图和[V?]标记的使用方法	科大讯飞
Rev1.1	2008-9-10	添加了章节 2.2、2.3，调整了 2.4 的内容，修改了 3.1 的内容	科大讯飞

目 录

1. 概述	- 4 -
1.1 简介	- 4 -
1.2 功能	- 4 -
1.3 特点	- 4 -
1.4 应用	- 5 -
2. XF-S4240 模块的使用说明	- 6 -
2.1 XF-S4240 模块的硬件接口	- 6 -
2.2 模块装配图和实物图	- 6 -
2.3 引脚信号定义	- 7 -
2.3.1 模块引脚排列信号图	- 7 -
2.3.2 模块引脚定义列表	- 7 -
2.4 XF-S4240 通讯方式	- 7 -
2.4.1 UART 通讯方式	- 8 -
2.4.2 SPI 通讯方式	- 8 -
2.4.3 I ² C 通讯方式	- 9 -
3. XF-S4240 语音合成板卡通讯协议	- 11 -
3.1 上位机发送的控制命令	- 11 -
3.2 板卡反馈信息	- 15 -
3.3 板卡工作状态引脚输出	- 16 -
3.4 上位机对 XF-S4240 板卡调用方式	- 16 -
3.4.1 简单调用	- 16 -
3.4.2 标准调用方式	- 16 -
3.5 协议发送示例程序	- 18 -
3.5.1 C 语言范例程序	- 18 -
3.5.2 汇编范例程序	- 19 -
4. 电路参考	- 21 -
4.1 复位方式	- 21 -
4.2 与上位机通讯连接	- 21 -
4.2.1 UART 通讯	- 21 -
4.2.2 SPI 通讯	- 22 -
4.2.3 I ² C 通讯	- 22 -
4.3 音频功放电路	- 22 -
5. 产品规格	- 24 -
5.1 电气特性	- 24 -
5.2 XF-S4240 外形尺寸	- 24 -
6. 输入文本的特殊标记	- 25 -
6.1 使用特殊控制标记	- 27 -
6.1.1 [b?] 标记的使用	- 27 -
6.1.2 [f?] 标记的使用	- 27 -

6.1.3	[i?] 标记的使用	- 27 -
6.1.4	[m?]标记的使用	- 27 -
6.1.5	[n?]标记的使用	- 27 -
6.1.6	[y?]标记的使用	- 27 -
6.1.7	[p?]标记的使用	- 28 -
6.1.8	[r?]标记的使用	- 28 -
6.1.9	[=?]标记的使用	- 28 -
6.1.10	[w]标记的使用	- 28 -
6.1.11	[x?]标记的使用	- 28 -
6.1.12	[s?]标记的使用	- 29 -
6.1.13	[t?]标记的使用	- 29 -
6.1.14	[v?]标记的使用	- 29 -
6.1.15	[d]标记的使用	- 29 -
7.	提示音使用范例	- 30 -
8.	附录	- 31 -
8.1	表格索引	- 31 -
8.2	图表索引	- 31 -
8.3	GB2312 编码简介	- 32 -
8.4	UNICODE 编码简介	- 32 -
8.5	GBK 编码简介	- 33 -
8.6	BIG5 编码简介	- 33 -

1. 概述

1.1 简介

XF-S4240 中文语音合成模块，是安徽科大讯飞信息科技股份有限公司（科大讯飞）推出的基于科大讯飞在嵌入式中文语音合成领域的最新研究成果——InterSound4.0 中文语音合成系统，而设计的一款中文语音合成模块。该模块可以通过异步串口（UART）、SPI 接口及 I²C 总线三种方式接收待合成的文本，直接合成为语音输出；主要是面向中高端应用，为其提供一套完整的物美价廉的语音解决方案。

1.2 功能

- 可合成任意的中文文本，支持英文字母的合成；
- 支持 GB2312、GBK、BIG5、UNICODE 四种内码格式的文本；
- 具有智能的文本分析处理算法，可正确的识别和处理数值、号码、时间日期及一些常用的度量衡符号，具备较强多音字处理和中文姓氏处理能力；
- 双发音人：男声、女声；
- 清晰、自然、准确的文语音合成效果；
- 集成提示音效，针对某些行业领域的常见语音提示音；
- 支持软件调节语速、语调、音量；
- 支持多种控制命令，包括：合成、停止、暂停合成、继续合成等；
- 支持多种文本控制标记，提升文本处理的正确率；
- 支持休眠功能，在休眠状态下可降低功耗；
- 支持 UART、SPI、I²C 三种数据通讯接口；
- Line out 音频输出；

1.3 特点

XF-S4240 中文语音合成模块是一款针对嵌入式应用领域而设计，具有合成任意中文文本的能力及英文字母合成的语音合成模块。该模块的主要特点是合成语音自然度高，控制接口简单方便，功能强大。

表格 1 XF-S4240 特点

项目	指标
文本内容	普通书面文本，每次合成的文本量最多可达 1024 字节
识别汉语拼音串	支持
字符集	GB2312 / GBK / BIG5 / UNICODE
提示音效	支持
语种	汉语普通话
中英文混读	英文按照字母发音
特殊符号识别	支持
多音字识别	支持普通多音字和姓氏多音字
发音人	男声、女声
发音风格	平铺直叙、一字一顿
语速调整	支持 10 级调整
语调调整	支持 10 级调整
音量调整	支持 10 级调整
工作状态	状态管脚输出工作状态、通过查询命令查询工作状态、自动返回工作状态（UART 接口）
文本输入接口	UART、SPI、I2C
声音输出	Line-out，接有源音箱即可播放声音

1.4 应用

- 车载 GPS 导航
- 车载电话、信息电话
- 公交车语音报站器
- 考勤机、打卡机
- 税控机、POS 机
- 智能仪器
- 智能玩具
- 排队机
- 自动售货机

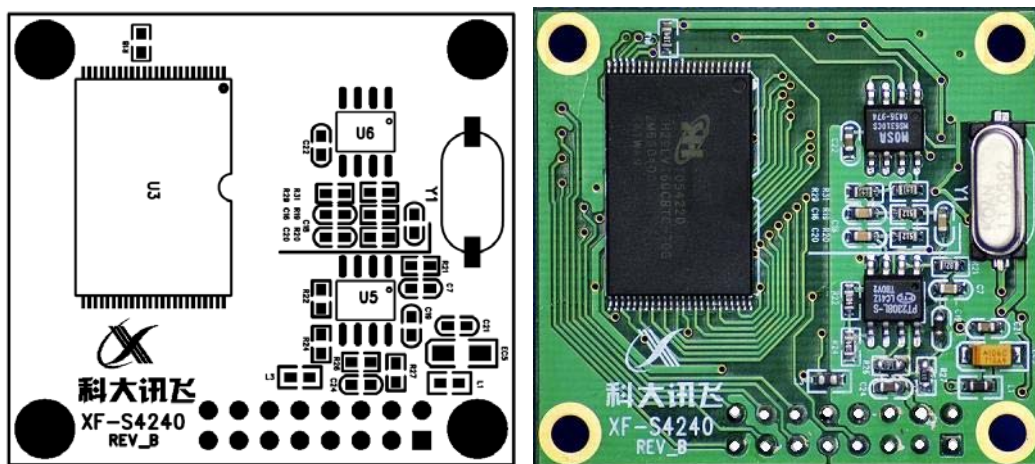
2. XF-S4240 模块的使用说明

正确的使用 XF-S4240 模块需要首先对模块的硬件接口进行了解。下面对模块的这些内容做一个介绍。

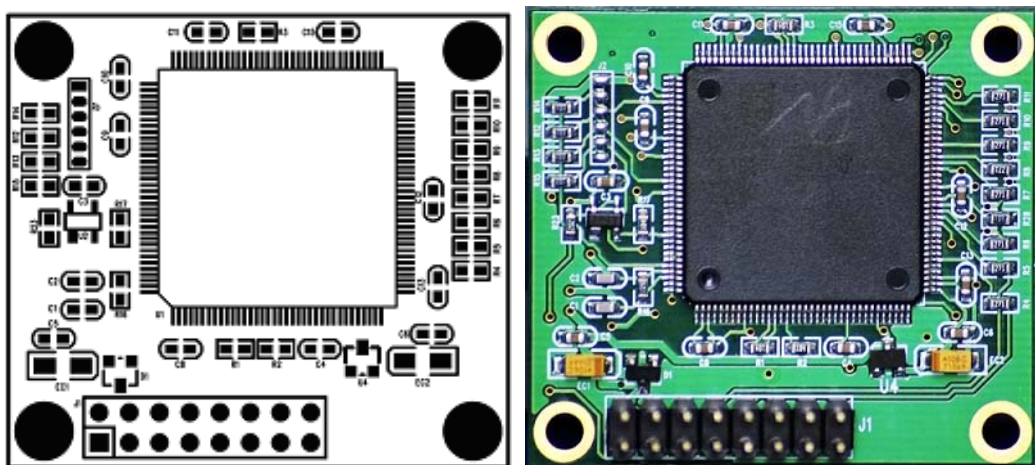
2.1 XF-S4240 模块的硬件接口

- 1) 工作电压：DC3.3V（范围 3.0V~3.6V）；
- 2) 提供通用异步串行数据通讯接口、SPI 数据通讯接口及 I²C 总线数据通讯接口；
- 3) 音频输出最大幅度 3.0V。

2.2 模块装配图和实物图



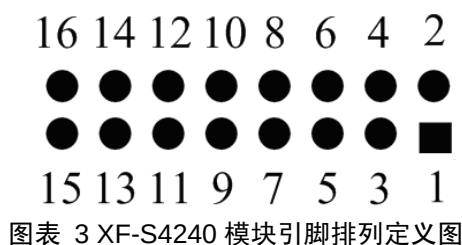
图表 1 XF-S4240 模块正面图



图表 2 XF-S4240 模块背面图

2.3 引脚信号定义

2.3.1 模块引脚排列信号图



图表 3 XF-S4240 模块引脚排列定义图

注意： XF-S4240 模块的硬件接口是 16 针的双排插针，在设计板卡的接口插座时，请注意与图表 3 中模块的接口保持一致。

2.3.2 模块引脚定义列表

表格 2 XF-S4240 引脚说明

引脚序号	引脚名称	功能描述	备注
1	Gnd	参考地信号	
2	Vcc	+3.3V 工作电源	
3	Gnd	参考地信号	
4	Line out	声音信号输出	
5	Gnd	参考地信号	
6	Gnd	参考地信号	
7	$\overline{RDY}/BSY$	工作状态指示输出	低电平：Ready 状态 高电平：Busy 状态
8	Rxd	UART 串行数据输入	
9	$\overline{RST}$	外部复位信号输入	低电平复位
10	Txd	UART 串行数据输出	
11	I2C_SCL	I2C 串行时钟输入	
12	I2C_SDA	I2C 串行数据接口	
13	SPI_SCK	SPI 串行时钟输入	
14	SPI_SSEL	SPI 使能选择输入	
15	SPI_MISO	SPI 串行数据输出	
16	SPI_MOSI	SPI 串行数据输入	

2.4 XF-S4240 通讯方式

XF-S4240 语音合成板卡支持 UART、SPI 和 I2C 三种通讯方式，用户上位机可选择其中任意一种通讯方式同合成板卡进行通讯，下面分别介绍板卡的这三种通讯方式。

2.4.1 UART 通讯方式



图表 4 通讯传输字节格式

- 1) 波特率: 9600 bps
- 2) 起始位: 1bit
- 3) 数据位: 8 bits
- 4) 停止位: 1 bit
- 5) 校验: 无

表格 3 XF-S4240 UART 通讯引脚描述

引脚序号	引脚名称	类型	描述
8	Rxd	输入	UART 串行数据输入
10	Txd	输出	UART 串行数据输出

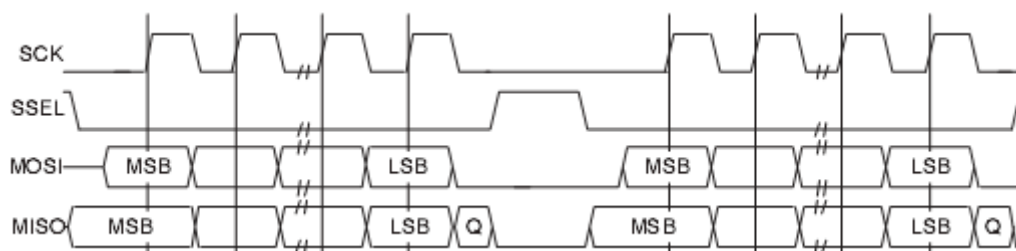
2.4.2 SPI 通讯方式

XF-S4240 的 SPI 接口是 4 线全双工同步串行通讯接口, 板卡在 SPI 通讯中设置为 Slave 身份, SPI 通讯所需的时钟信号由上位机提供, 即上位机作为 SPI 通讯中的 Master 身份。传输数据的位数为 8bits。

表格 4 XF-S4240 SPI 通讯引脚描述

引脚序号	引脚名称	类型	描述
13	SPI_SCK	输入	SPI 同步时钟输入
14	SPI_SSEL	输入	SPI 使能选择输入
15	SPI_MISO	输出	SPI 串行数据输出
16	SPI_MOSI	输入	SPI 串行数据输入

下图是 SPI 通讯的时序图, SPI 通讯属于同步串行通讯, 需要 Master 端提供同步时钟信号。在和板卡的通讯中, 上位机将提供 SCK 同步时钟信号, 在 SCK 的上升沿板卡和上位机同时锁存 1bit 数据, 每传输 8bits 数据完成一个字节数据的传输。SSEL 信号的功能是选择此 XF-S4240 板卡位上位机 SPI 通讯的使能 Slave, 在传输过程中, SPI 使能选择信号 SSEL 要保持低电平, 不然无法正常通讯。



图表 5 SPI 通讯时序

注意：使用 SSEL 选择使能信号是为了方便用户在 SPI 通讯接口上挂接多个 XF-S4240 板卡（或其他 SPI 接口的设备），在和选定的 XF-S4240 板卡通讯时要将此设备的 SSEL 置为低电平。

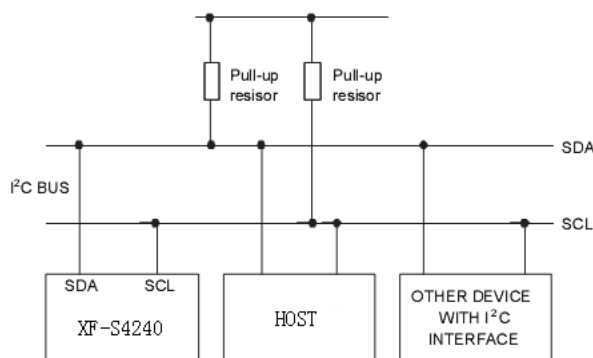
2.4.3 I²C 通讯方式

XF-S4240 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。

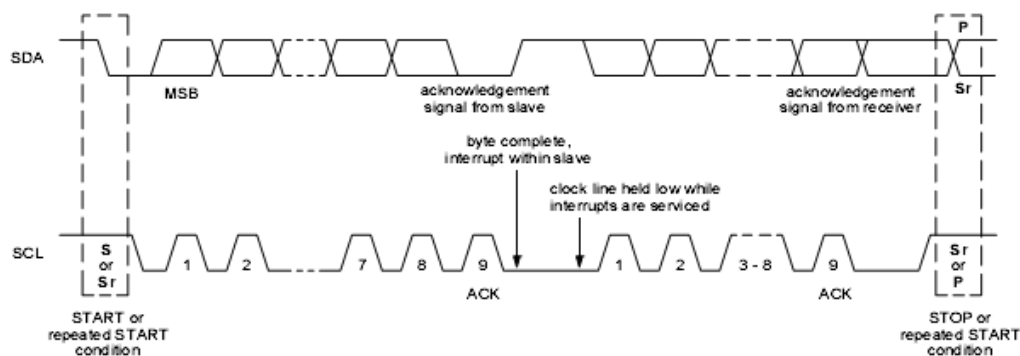
表格 5 I²C 总线通讯引脚描述

引脚序号	引脚名称	类型	描述
11	I2C_SCL	输入	I ² C 同步时钟输入
12	I2C_SDA	输入/输出	I ² C 串行数据接口

挂接在 I²C 总线上的设备只使用两条线（SCL 和 SDA）来实现双向通讯。I²C 总线上能挂接多个设备，由于使用 7bits 数据作为设备地址，所以总线上最多能挂接 127 个设备。XF-S4240 板卡在 I²C 总线上的连接方式如图所示。


图表 6 XF-S4240 I²C 总线挂接

I²C 总线通讯的时序如图所示，传输的数据位数是 8bits，每次传输数据的字节数无限制，每一个字节传输后返回确认位。



图表 7 I²C 总线时序图

通过 I²C 总线，用户可以方便的同时使用多个 XF-S4240 板卡，最多 127 个。每个板卡都有自己的地址，用户向不同地址的板卡发送文本即可实现一个上位机同时操作多个板卡进行不同的语音合成。

XF-S4240 板卡作为 I2C 接口的从设备，默认的设备地址为 **0x80**。对于需要挂接多个 XF-S4240 板卡的系统，我们提供更改设备地址的服务。

XF-S4240 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](#)”。

3. XF-S4240 语音合成板卡通讯协议

3.1 上位机发送的控制命令

上位机发送给 XF-S4240 的所有命令和数据都需要用“帧”的方式进行封装后传输，如表 1 所示。帧结构由帧头标志、数据区长度和数据区三部分组成。表 2 是对命令帧格式の説明。

表格 6 命令帧封装格式

帧头	数据区长度	数据区
0xFD(1Byte)	0xXX, 0xXX (2Byte)	Data

表格 7 命令帧说明

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 1024Bytes	命令字和命令参数，长度和“数据区长度”一致

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板卡的各种功能，表格 8 中详细描述了数据区的命令字和命令参数。

表格 8 数据区命令字和命令参数

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	TTS 系统状态查询命令
	0x88	系统进入 Power Down 模式，Reset 之后恢复
命令参数	不同命令字有不同参数列表，详见各命令字说明	

1. 语音合成命令

表格 9 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	<table><tr><td>帧头</td><td colspan="2">数据区长度</td><td colspan="3">数据区</td></tr><tr><td rowspan="2">0xFD</td><td>高字节</td><td>低字节</td><td>命令字</td><td>文本编码格式</td><td>待合成文本</td></tr><tr><td>0xHH</td><td>0xLL</td><td>0x01</td><td>0x00~0x03</td><td>.... ..</td></tr></table>					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	
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 XF-S4240 模块正在合成文本的时候，如果又接收到一帧有效的数据，模块会立即停止当前正在合成的文本，转而合成所接收到的最新的文本																					

2. 停止合成命令

表格 10 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
		0xHH	0xLL	0x02
示例	停止合成			
	0xFD	0x00	0x01	0x02

3. 暂停合成命令

表格 11 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
		0xHH	0xLL	0x03
示例	暂停合成			
	0xFD	0x00	0x01	0x03

4. 恢复合成命令

表格 12 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
		0xHH	0xLL	0x04
示例	恢复合成			
	0xFD	0x00	0x01	0x04

5. 系统状态查询命令

表格 13 系统状态查询命令

名称	发送的数据	说明		
命令字	0x21	通过该命令来判断 TTS 模块是否正常工作，以及获取相应参数，返回 0x4E 表明系统仍在合成中，返回 0x4F 表明系统处于空闲状态		
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
		0xHH	0xLL	0x21
示例	状态查询			
	0xFD	0x00	0x01	0x21

6. 进入 Power Down 状态命令

表格 14 进入休眠状态命令

名称	发送的数据	说明		
命令字	0x88	进入 POWER DOWN 状态命令，RST 后恢复		
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
		0xHH	0xLL	0x88
示例	进入 Power Down 状态			
	0xFD	0x00	0x01	0x88

3.2 板卡反馈信息

XF-S4240 在初始化成功时会回传一个字节的“初始化成功”信息，初始化不成功上位机收不到此信息。在收到命令帧头“0xFD”后板卡会判断此命令帧正确与否，如果命令帧正确则返回“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。在板卡收到 TTS’系统状态查询命令时，如果 TTS 系统处于合成工作状态则返回“TTS 系统忙碌”信息，如果 TTS 系统处于空闲状态则返回“TTS 系统空闲”信息。在一帧数据合成完毕后，TTS 系统会自动返回一次“TTS 系统空闲”信息。

表格 15 板卡反馈信息

名称	发送的数据	触发条件
初始化成功回传	0x4A	系统初始化成功
收到正确的命令帧回传	0x41	收到正确的命令帧
收到错误命令帧回传	0x45	收到错误的命令帧
TTS 系统空闲	0x4F	1. 当一帧数据合成完以后，TTS 系统空闲状态回传 0x4F 2. 收到“状态查询命令”，TTS 系统处于空闲状态回传 0x4F
TTS 系统忙碌	0x4E	1. 收到“状态查询命令”，TTS 系统非空闲状态回传 0x4E

3.3 板卡工作状态引脚输出

表格 16 工作状态引脚描述

引脚序号	引脚名称	类型	描述
7	$\overline{\text{Rdy/Bsy}}$	输出	低电平: Ready 状态 高电平: Busy 状态

XF-S4240 板卡提供了一个硬件引脚指示板卡的工作状态,用户可以通过查询此状态输出引脚的电平,来判断板卡工作的状态。当此引脚处于低电平时,表明板卡处于空闲状态,没有合成文本;当引脚处于高电平状态时,表明板卡处于合成文本的工作状态。

3.4 上位机对 XF-S4240 板卡调用方式

3.4.1 简单调用

简单调用针对初级用户或应用比较简单的情况。用户不用关心板卡的工作状态,只需要发送文本,板卡会将接收的文本合成为语音输出。

在简单调用情况下,上位机只要与板卡之间建立起 UART、SPI 或者 I²C 三种通信方式中的任意一种,即可发送合成命令来实现文本的合成,上位机不需要理睬板卡的反馈信息和状态输出, XF-S4240 立即输出合成的语音。

提示: 如果前一帧文本还没有合成完毕,再发送文本到板卡就会打断前次合成,而执行新的合成。

3.4.2 标准调用方式

对于一般情况,上位机需要了解板卡的工作状态,以更精确的控制板卡的动作:比如需要确保上次文本被完整合成之后,再合成下一段文本。为了实现这个目的,有三种实现方法:

1. 通过查询 $\overline{\text{Rdy/Bsy}}$ 引脚来了解板卡的工作状态,如此引脚处于高电平状态,说明板卡正在处于文本合成状态,需要继续等待,直到此引脚变成低电平状态后,系统处于空闲状态,再给板卡发送下一段文本进行合成;
2. 向板卡发送系统状态查询命令,通过返回值即可知道系统是否空闲(详见表格 13);
3. 接收来自板卡的反馈信息,当文本合成完毕后,板卡会反馈“0x4F”字符,表示系统已经空闲(对于 UART 通讯方式上位机可以直接接收此信息,对于 SPI 和 I²C 通讯方式,由于板卡是 Slave 身份,需要上位机主动读取此信息);

应用举例如下:假设需要合成的文本为 2000 字节,超过了板卡一个命令帧所能容纳的最大文本长度(1024 字节),这时分两次给板卡发送文本信息。程序过程如下:

- 1、上位机先给板卡发送一个文本合成命令帧,携带 1000 个字节的文本;
- 2、上位机等待板卡返回播放完毕回传信息,直到收到板卡反馈“0x4F”系统空闲信息;

- 3、上位机给板卡再次发送一个文本合成命令帧，发送出剩下的 1000 个字节文本信息。

3.5 协议发送示例程序

3.5.1 C 语言范例程序

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：“欢迎使用科大讯飞语音合成板卡。”。

```
void    main(void)
{
    /*****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*****/
    //需要发送的文本
    unsigned char code text[ ]={"欢迎使用科大讯飞语音合成板卡。"};
    unsigned int length = 0;
    int i = 0;
    p = text;
    length = strlen(p); //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 = length + 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<length; i++) //依次发送待合成的文本数据
{
    SBUF = *p;
    While (TI == 0) wait ();
    TI = 0;
    p++;
}
}
```

3.5.2 汇编范例程序

下面是 51 单片机为控制上位机的汇编语言示例，演示向板卡发送“科大讯飞。”这段文本进行合成：

“科大讯飞。”的 GB 码是：

“科” 0xBF C6 “大” 0xB4 F3 “讯” 0xD1 B6 “飞” 0xB7 C9 “。” 0xA1 A3

```
Send_TTS_Command:
    Mov sbuf, #FDh      ;帧头标志
    Mov sbuf, #00h      ;发送合成文本长度，高字节在前，低字节在后
    Mov sbuf, #0Ch      ;合成文本长度的低字节
    Mov sbuf, #01h      ;命令字
    Mov sbuf, #00h      ;文本编码格式
    Mov sbuf, #BFh      ;发送“科”高8位
    Mov sbuf, #C6h      ;发送“科”低8位
    Mov sbuf, #B4h      ;发送“大”高8位
    Mov sbuf, #F3h      ;发送“大”低8位
    Mov sbuf, #D1h      ;发送“讯”高8位
    Mov sbuf, #B6h      ;发送“讯”低8位
    Mov sbuf, #B7h      ;发送“飞”高8位
    Mov sbuf, #C9h      ;发送“飞”低8位
    Mov sbuf, #A1h      ;发送“。”高8位
    Mov sbuf, #A3h      ;发送“。”低8位
```

以上为发送一帧 TTS 文本数据的程序模块

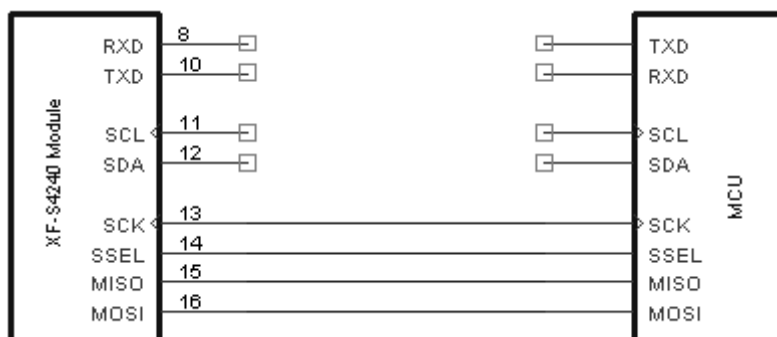
发送后可接收到板卡的反馈信号，如果是“41 4F”，则说明文本被正确接收，且表明合成完毕，板卡处于空闲状态；如果收到的是“45”，则说明文本没有正确的收到或是合成，则需要重新发送或者复位。

注：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

进行判断当前板卡的工作状态，然后才能发送下一个数据。

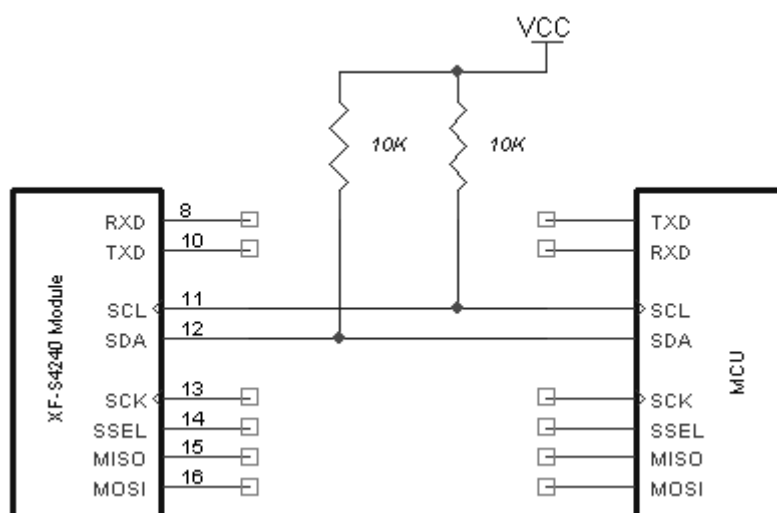
第 - 21 - 页

4.2.2 SPI 通讯



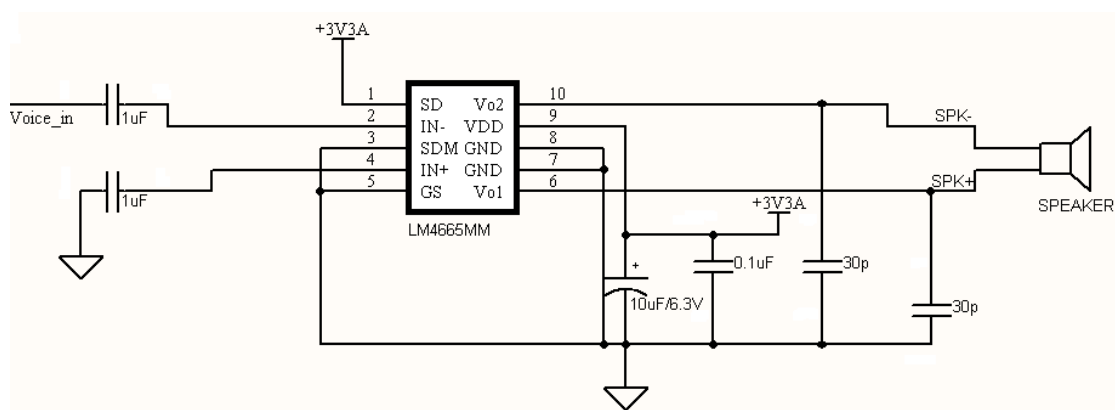
图表 10 XF-S4240 模块与 MCU 的 SPI 接口连接示意图

4.2.3 I²C 通讯



图表 11 XF-S4240 模块与 MCU 的 I2C 接口连接示意图

4.3 音频功放电路



图表 12 XF-S4240 模块与音频功放 LM4665 的接口示意图

5. 产品规格

5.1 电气特性

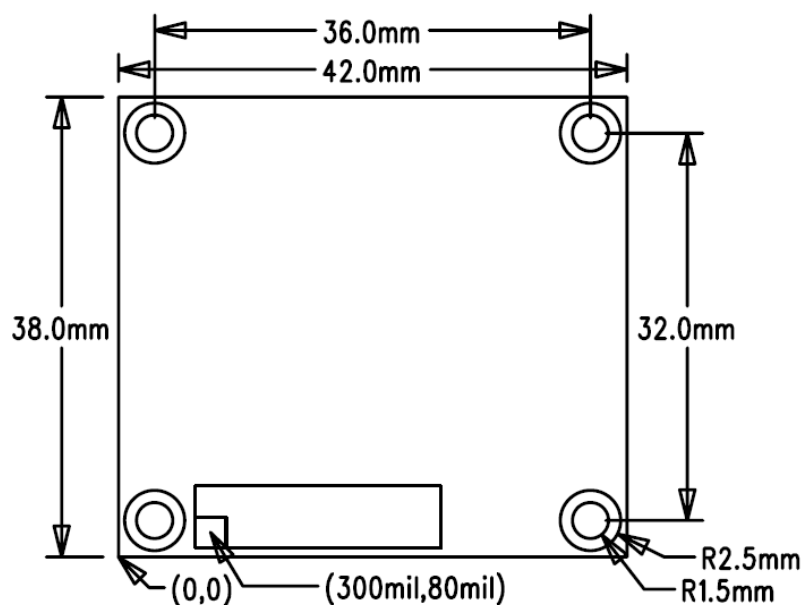
表格 17 极限值

参 数	符 号	最小值	最大值
电源电压	V_{DD}	-0.5	3.6V
引脚输入电压范围	V_{IN}	-0.5	4.5V
工作温度	T_A	-25℃	+75℃
储存温度	T_{STO}	-45℃	+100℃

表格 18 静态特性

参 数	符 号	参 考 值			单 位
		最小	典型	最大	
工作电压	V_{CC}	3.0	3.3	3.6	V
工作电流	I_{OP}	--	45	50	mA
输入高电平	V_{IH}	2	--	--	V
输入低电平	V_{IL}	--	--	0.8	V
输出电流	I_{OH}	--	-4	--	mA
吸收电流	I_{OL}	--	4	--	mA

5.2 XF-S4240 外形尺寸



图表 13 XF-S4240 尺寸图

6. 输入文本的特殊标记

XF-S4240 板卡基本上能正确的识别汉语中的多音字以及能正确判断人名和称谓中的姓氏读法。系统可以识别时间、金额、比分等特殊的文本格式，并按人们的习惯方式朗读。在数值、号码等文本的合成效果上也有较好的表现。但是由于汉语博大精深，系统不可能做到全部正确识别，为了达到百分之百的合成效果，系统为用户提供了多种调节功能，使得合成效果有更丰富的表现。只需在文本中插入参数调节标识，即可进行参数调节，使用起来非常直观和方便。

用户可以根据自己的需求利用标记的方法改变输出的声音大小，实现对话速调节的支持，实现对话调调节的支持，选择不同的发音人（女声/男声），是否朗读标点符号，并针对多音字、数字、姓氏等等读法有调节标记，使得用户只需在文本中插入标记即可实现对模块控制的功能，无需发送多余命令，极大的方便了用户使用。下面详细介绍系统的各种标记：

表格 19 文本特殊标记说明

作用	控制标识	详细说明	系统默认
标点符号读法	[b?]	? 为 0, 标点符号按正常逻辑作停顿处理	[b0]
		? 为 1, 读出标点符号的名称	
设置发音风格	[f?]	? 为 0, 一字一顿的风格	[f1]
		? 为 1, 叙述的风格	
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?]	? 为 3, 设置发音人为女声	[m3]
		? 为 4, 设置发音人为男声	
设置数字处理策略	[n?]	? 为 0, 自动判断	[n0]
		? 为 1, 数字作号码处理	
		? 为 2, 数字作数值处理	
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, 自动判断姓名读音	[r0]
		? 为 1, 强制使用姓名读音规则	
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]
强制分词	[w]	强制分词, 防止后接短语中出现换气停顿	
设置号码中“1”的读法	[y?]	? 为 0, 合成号码“1”时读成“幺”	[y0]
		? 为 1, 合成号码“1”时读成“一”	
设置提示音处理策略	[x?]	? 为 0, 不使用提示音	[x1]
		? 为 1, 使用提示音	
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]	所有设置恢复为默认值	

注:

- 1)、所有的控制标识均为半角字符。
- 2)、控制标识为全局控制标识, 也就是只要用了一次, 在不对模块复位或者断电的条件下, 其后发送给模块的所有文本都会处于它的控制之下, 除非用相应的 [d]恢复默认设置。
- 4)、如果[x0]、[x1]标记或是[r0]、[r1]标记在同一句语句中使用那么需要在[x0]或[x1]前加符号“,”。例如: “[x0]来电号码,[x1]来电号码”; “[r1]解小东, [r0]解放军”。
- 5)、当模块掉电或是复位后, 模块将恢复到所有的默认值, 原来的设置过的标识失去了作用, 需要重新设置。

6.1 使用特殊控制标记

6.1.1 [b?] 标记的使用

示例文本	模块解释
[b0]科大, 讯飞	系统合成：“科大”（停顿）“讯飞”
[b1]科大, 讯飞	系统合成：“科大逗号讯飞”

6.1.2 [f?] 标记的使用

示例文本	模块解释
[f0]科大讯飞	系统按照一字一顿的风格合成
[f1]科大讯飞	系统按照叙述的风格

6.1.3 [i?] 标记的使用

示例文本	模块解释
[i0]科大讯 fei1	系统不识别拼音格式。合成为：“科大讯 FEI 一”
[i1]科大讯 fei1	系统识别拼音格式。合成为：“科大讯飞”

6.1.4 [m?]标记的使用

示例文本	模块解释
[m3]科大讯飞	女声合成：“科大讯飞”
[m4]科大讯飞	男声合成：“科大讯飞”

6.1.5 [n?]标记的使用

示例文本	模块解释
[n0]234343545	系统自动判断。读作：二亿三千四百三十四万三千五百四十五
[n1]234343545	系统强制按照数值的方式合成数字串。读作：二三四三四三五四五
[n2]234343545	系统强制按照号码的方式合成数字串。读作：二亿三千四百三十四万三千五百四十五

6.1.6 [y?]标记的使用

示例文本	模块解释
[ls]2[y0]134856[]	系统按照“幺”的读法合成号码文本中的“一”。读作：二幺三四八五六
[ls]2[y1]134856[]	系统按照“一”的读法合成号码文本中的“一”。读

	作：二一三四八五六
--	-----------

*注意：此标记必须是在合成号码类型文本的时候才有效；

6.1.7 [p?]标记的使用

示例文本	模块解释
欢迎使用[p500]科大讯飞[p1000]语音合成芯片	系统在“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 8 秒。

6.1.8 [r?]标记的使用

示例文本	模块解释
[r0]曾国藩	系统自动的识别多音字，按照默认的方式合成。读作：céngguófān
[r1]曾国藩	系统强制的将文本的第一个字按照姓名的方式合成。读作：zēngguófān

一般用于句首在同一文本中可用该标记进行模式切换。

例子：参加人员有：[r1]解小东、曾国庆、朴志全等[r0]解放军指战员

6.1.9 [=?]标记的使用

示例文本	模块解释
曾[=zeng1]国藩	系统强制的将“曾”按照“zēng”音合成播报

6.1.10 [w]标记的使用

示例文本	模块解释
科[w]大讯飞	系统强制将“科”和后面的“大讯飞”分开合成。

6.1.11 [x?]标记的使用

1) . 正确用法

示例文本	模块解释
预先制定的提示音铃声 [x0]sounda 为，[x1]sounda	[x0]和[x1]均起作用。读作：预先制定的提示音铃声 sounda 为（提示音 sounda 的铃声）

2) . 错误用法，

示例文本	模块解释
预先制定的提示音铃声 [x0]sounda 铃声为[x1]sounda	只有[x1]起作用。读作：预先制定的提示音铃声（提示音 sounda 的铃声）为（提示音 sounda 的铃声）

6.1.12 [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	软件语速的调节功能, 提供 10 级的语速调节, 系统默认的语速大小为 5, 最小为 0, 最大为 10。所以本句中“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。

6.1.13 [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	软件语调的调节功能, 提供 10 级的语调调节, 系统默认的语调大小为 5, 最小为 0, 最大为 10。所以本句中“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。

6.1.14 [v?]标记的使用

示例文本	模块解释
欢迎[v1]使用[v5]科大讯飞[v9]语音合成芯片	软件音量的调节功能, 提供 10 级的音量调节, 系统默认的音量大小为 8, 最小为 0 (此时相当于静音), 最大为 9。所以本句中“欢迎”按照 8 级音量合成, “使用”为 1 级音量, “科大讯飞”为 5 级音量, “语音合成芯片”为 9 级音量。

6.1.15 [d]标记的使用

示例文本	模块解释
[r1]曾国藩, [d], 曾国藩	系统首先将文本的第一个字按照姓名的方式合成, 经过[d]标记的恢复, 回到自动识别的状态。读作: zēng guó fān, céng guó fān

7. 提示音使用范例

系统中还提供了其他的一些简单的声音提示音，这些都可以广泛的使用在一些公共信息场合或是需要提醒的场合。下面列表中是目前系统中含有的音乐提示音内容及建议用途：

表格 20 提示音列表

提示音文本	sounda	消息提示音
	soundb	消息提示音
	soundd	电话铃音
	soundg	门铃音
	soundi	消息提示音

提示音在使用上没有特殊性，仍然是按照普通的文本方式发送，但是需要注意的是**提示音文本需要能够单独成为一个分句**，系统就会自动识别。下面介绍几个使用实例：

“**soundd**，您好，你有来电，来电号码 05511860。”

“**soundd**。您好，你有来电，来电号码 05511860。”

“**soundd** | 您好，你有来电，来电号码 05511860。”

例子说明：此句为提示音和自然文本集合的一个文本，蓝色字体为提示音，他们只要按照普通文本的方式发送即可，但是必须要成为一个单独的文本，所以在提示音和提示音之间或是提示音和自然文本之间增加了一个间隔符“，”，间隔符主要就是起到分隔的作用使得文本成为单独的一个分句交给系统。除了“，”、“。”、“|”外，“！”、“？”等符号也是间隔符，但是“，”、“。”、“！”、“？”符号都是带有停顿的，所以在有些情况下会造成合成效果的不流畅，而“|”只有间隔的作用没有停顿的作用。用户可以根据自己的需求来使用相应的间隔符。

8. 附录

8.1 表格索引

表格 1 XF-S4240 特点	- 5 -
表格 2 XF-S4240 引脚说明	- 7 -
表格 3 XF-S4240 UART 通讯引脚描述	- 8 -
表格 4 XF-S4240 SPI 通讯引脚描述	- 8 -
表格 5 I ² C 总线通讯引脚描述	- 9 -
表格 6 命令帧封装格式	- 11 -
表格 7 命令帧说明	- 11 -
表格 8 数据区命令字和命令参数	- 11 -
表格 9 语音合成命令	- 12 -
表格 10 停止合成命令	- 13 -
表格 11 暂停合成命令	- 13 -
表格 12 恢复合成命令	- 14 -
表格 13 系统状态查询命令	- 14 -
表格 14 进入休眠状态命令	- 15 -
表格 15 板卡反馈信息	- 15 -
表格 16 工作状态引脚描述	- 16 -
表格 17 极限值	- 24 -
表格 18 静态特性	- 24 -
表格 19 文本特殊标记说明	- 26 -
表格 20 提示音列表	- 30 -

8.2 图表索引

图表 1 XF-S4240 模块正面图	- 6 -
图表 2 XF-S4240 模块背面图	- 6 -
图表 3 XF-S4240 模块引脚排列定义图	- 7 -
图表 4 通讯传输字节格式	- 8 -
图表 5 SPI 通讯时序	- 9 -
图表 6 XF-S4240 I ² C 总线挂接	- 9 -
图表 7 I ² C 总线时序图	- 10 -
图表 8 XF-S4240 模块与 PC 的异步串口连接示意图	- 21 -
图表 9 XF-S4240 模块与 MCU 的异步串口连接示意图	- 21 -
图表 10 XF-S4240 模块与 MCU 的 SPI 接口连接示意图	- 22 -
图表 11 XF-S4240 模块与 MCU 的 I2C 接口连接示意图	- 22 -
图表 12 XF-S4240 模块与音频功放 LM4665 的接口示意图	- 23 -
图表 13 XF-S4240 尺寸图	- 24 -

8.3 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”。

8.4 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，仅欧洲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工具的存在，是近来全球软件技术最重要的发展趋势。

8.5 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的承上启下的作用。

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 Unicode2.0。

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。

8.6 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。