

# 软件开发手册

版本: V1.0

# 目录

<b>1. 系统功能说明</b> .....	<b>3</b>
1.1 功能特性 .....	3
1.2 通信方式 .....	3
<b>2. 通信协议</b> .....	<b>4</b>
2.1 命令帧封装格式 .....	4
2.2 命令帧定义 .....	5
2.3 系统命令 .....	5
2.4 语音合成功能支持的控制命令 .....	6
2.5 文本缓存功能的控制命令 .....	8
2.6 上传协议介绍 .....	10
2.7 上传状态帧 .....	10
<b>3. 文本标注方法</b> .....	<b>11</b>

# 1. 系统功能说明

## 1.1 功能特性

◇ 支持纯中文、常规英文文本（如字母、单词、短句）以及中英文混合合成  
可实现多音字、生僻字、数字、数值、日期、时间、字母正确流畅的合成，播放清晰无杂音，支持实时更新的特殊播报（如：字母、数字穿插播报、大写字母缩写）。

◇ 支持五种文本编码方式

当前方案支持 GB2312、GBK、BIG5、UTF16LE 和 UTF8 五种编码方式。每次合成的文本量最多支持 4K 字节（经过相应的解码格式解码后），（中文不超过 2000 个字，具体跟编解码格式相关）。具体的每种解码格式字符占用空间参考表 1-1。

解码格式	英文字符	中文	英文标点	中文标点
GB2312	1 字节	2 字节	1 字节	2 字节
GBK	1 字节	2 字节	1 字节	2 字节
BIG5	1 字节	2 字节	1 字节	2 字节
UTF16LE	2 字节	2 字节	2 字节	2 字节
UTF8	1 字节	3 字节	1 字节	3 字节

表 1-1 不同解码方式字符占用空间

◇ 支持多种控制命令

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、进入省电模式、唤醒等控制命令。控制器通过通信接口发送控制命令可以对芯片进行相应的控制。

◇ 可查询芯片的工作状态

通过读芯片自动返回的工作状态字、发送查询命令可以获得芯片工作状态的回传数据。

## 1.2 通信方式

芯片和上位机之间都是通过下传命令和状态反馈依次交替的方式进行通信。下传命令是由上位机发起的。状态反馈是由芯片发起的；状态即通信状态，包含帧标记，帧数据长度错误，命令字错误，命令正确执行等。上位机向芯片下传命令后，必然会收到芯片的状态反馈；上位机只有在收到芯片的状态反馈后，才能进行下一次的下传命令操作。

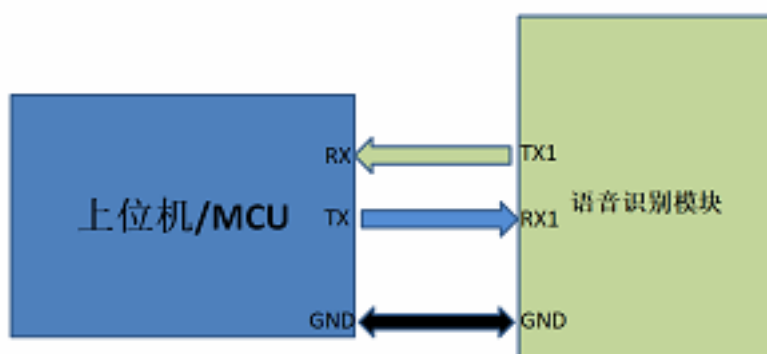
芯片在收到上位机的下传命令后，对命令的解析是非阻塞操作，可以较快的反馈状

态给上位机。

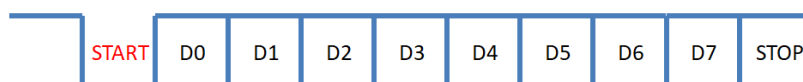
对于 UART 通信接口，上位机发送下传命令后，芯片会给上位机反馈状态。上位机在收到反馈状态后，再进行下一次的命令下传。

注意：芯片在上电后并系统初始化完成后，会向上位机反馈系统 Ready 的状态。因此对于 UART 通信接口，上位机必须先通过读操作来获取 Ready 状态后，才能进行第一次的命令下传。

#### ■ 硬件连接:



#### ■ 通信传输字节格式



波特率:115200bps

起始位:1 bit

数据位:8 bits

停止位:1 bit

校验位:无

## 2. 通信协议

平台通信协议是指上位机和芯片之间的通信协议，即上位机如何控制模块，模块如何将信息反馈给上位机。通信协议分为下传协议和上传协议两部分。

### 2.1 命令帧封装格式

上位机通过下传命令帧来控制芯片。下传命令帧的格式定义就是下传协议。上位机发

送给芯片的所有控制命令都需要用“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	0xXX, 0xXX	Data

## 2.2 命令帧定义

下传命令帧分为 2 部分：帧头，帧数据。其中帧头分为 2 部分：标记，帧长度。帧数据包含：命令字，命令参数（可选），命令数据。具体说明如下：

下传命令帧				
帧头		帧数据		
标记	帧数据长度	命令字	命令参数	命令数据
“0xFD” (1 字节)	“0xHH, 0xLL” (2 字节)	“命令” (1 字节)	“0xXX (可选)” (1 字节)	数据 (N 字节)

具体说明如下：

名称	长度	说明
标记	1 字节	由一个字节表示，固定为十六进制的“0xFD”
帧数据长度	2 字节	由两个字节表示，高字节在前，低字节在后
命令字	1 字节	由一个字节表示，表示具体的下传命令
命令参数	1 字节	由一个字节表示，此参数可选
命令数据	N 字节	表示具体的文本数据

## 2.3 系统命令

### ◇ 状态查询命令

名称	发送的数据	说明
命令字	0x21	查询当前系统工作运行状态。 模块在收到该命令后，将反馈运行状态。运行状态分为 Idle 和 Busy 两种。
参数列表	无	-
命令帧格式结构	标记	帧数据长度
	0xFD	0x00      0x01
		命令字
		0x21

### ◇ 休眠命令

名称	发送的数据	说明
命令字	0x88	系统进入到休眠状态。

		下位机收到此命令后, 会停止 TTS, 然后系统进入低功耗模式, 状态为 idle, 并返回 ok		
参数列表	无	-		
命令帧格式结构	标记	帧数据长度		命令字
	0xFD	0x00	0x01	0x88

#### ◇ 退出休眠命令

名称	发送的数据	说明		
命令字	0xFF	系统从休眠状态退出, 并把 TTS 准备好。		
参数列表	无			
命令帧格式结构	标记	帧数据长度		命令字
	0xFD	0x00	0x01	0xFF

注意：芯片进入省电模式后, 需要发送唤醒命令到芯片的通信接口上, 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。

## 2.4 语音合成功能支持的控制命令

下传命令包是指上位机发送给芯片所有下传命令帧, 如下:

命令	命令字	说明
开始命令	0x01	启动 TTS 开始合成此命令帧的 TEXT, PAGE 为文本的编码方式
停止命令	0x02	停止 TTS 的合成
暂停命令	0x03	暂停 TTS 的合成
继续命令	0x04	从暂停的位置开始继续合成
查询命令	0x21	查询芯片的状态, 返回状态只有 busy 和 idle 两种
cache 命令	0x31	把此命令帧的 TEXT 存储到 C4201-T00C 的第 INDEX 个缓存块开始的缓存里面。INDEX 的范围为 0~15, 每个缓存的大小是 256Byte, 因此 TEXT 的最大长度 4096 个字节
cplay 命令	0x32	开始合成芯片缓存里面的文本, 注意是从第 0 个缓存开始合成, 直到文本结束。REPEAT 为播放次数 (1~15), PAGE 为文本编码方式
睡眠命令	0x88	此命令会使芯片停止 TTS 合成, 进入低功耗模式, 此时状态为 idle, 并返回 ok。低功耗模式下可以用查询命令来获取状态。任何上位机命令都会使芯片退出低功耗模式 (除了查询命令和睡眠命令), 但是不会执行命令。退出低功耗模式后, 才能正常的解析和执行命令
唤醒命令	0xFF	使芯片退出低功耗模式。接收到此命令后, 芯片会退出低功耗模式, 并使 TTS 准备就绪, 并返回 ok

具体说明如下：

命令名称	下传命令帧					
	帧头			帧数据		
	TAG	FLENH	FLENL	CMD	AGRV	TEXT
开始命令 (start)	0xFD	0x00	LL	0x01	PAGE	TEXT
停止命令 (stop)	0xFD	0x00	0x01	0x02		
暂停命令 (pause)	0xFD	0x00	0x01	0x03		
继续命令 (resume)	0xFD	0x00	0x01	0x04		
查询命令 (query)	0xFD	0x00	0x01	0x21		
cache 命令 (cache)	0xFD	0x00	LL	0x31	INDEX	TEXT
cplay 命令 (cplay)	0xFD	0x00	0x02	0x32	REPEAT_PAGE	
睡眠命令 (sleep)	0xFD	0x00	0x01	0x88		
唤醒命令 (wakeup)	0xFD	0x00	0x01	0xFF		

#### ◇ 开始命令

名称	发送的数据	说明
命令字	0x01	启动 TTS 合成文本。 下位机开始启动 TTS，合成此命令帧的文本，并进行播音。下位机收到此命令后，会停止 TTS，然后系统进入低功耗模式，状态为 idle，并返回 ok
参数列表	[PAGE] 代码页	1 字节，支持如下代码页： 【0x00】GB2312； 【0x01】GBK； 【0x02】BIG5； 【0x03】UTF16LE； 【0x04】UTF8；
	[TEXT]合成文本	待合成文本，最大支持 4000 个字节（2000 个字）文本。根据数据帧长度以及代码页类型对文本进行解析。

命令帧格式结构	标记	帧数据长度		命令字	参数	文本
	0xFD	0x00	0xLL	0x01	PAGE	TEXT (N 字节)
UART 示例	例如，合成不同编码格式的“语音合成”： 合成 GB2312 格式“语音合成”(0xD3, 0xEF, 0xD2, 0xF4, 0xBA, 0xCF, 0xB3, 0xC9)：					
	0xFD	0x00	0x0A	0x01	0x00	0xD3 0xEF 0xD2
	0xF4	0xBA	0xCF	0xB3	0xC9	

◇ 停止命令

名称	发送的数据		说明	
命令字	0x02		停止 TTS 合成。	
参数列表	无			
命令帧格式结构	标记	帧数据长度		命令字
	0xFD	0x00	0x01	0x02

◇ 暂停命令

名称	发送的数据		说明	
命令字	0x03		暂停 TTS 合成。 注意：暂停 TTS 不是停止 TTS。前者可从暂停的位置继续播放，后者则需重新启动 TTS 来合成。	
参数列表	无			
命令帧格式结构	标记	帧数据长度		命令字
	0xFD	0x00	0x01	0x03

◇ 继续命令

名称	发送的数据		说明	
命令字	0x04		继续从暂停的位置，进行 TTS 合成。	
参数列表	无			
命令帧格式结构	标记	帧数据长度		命令字
	0xFD	0x00	0x01	0x04

## 2.5 文本缓存功能的控制命令



◇ Cache 命令

名称	发送的数据	说明						
命令字	0x31	把文本数据缓存到下位机的缓存里面。						
参数列表	[INDEX] 缓存块索引	1 字节, 范围为 0-15。表示把数据缓存到从下位机的第 INDEX 块缓存开始的位置, 每个缓存大小 256 字节, 共 16 块, 即 4096 字节						
	[TEXT] 合成文本	待合成文本, 最大支持 4000 个字节 (2000 个字) 文本。						
命令帧格式结构	标记	帧数据长度		命令字	参数	文本		
	0xFD	0x00	0xLL	0x31	INDEX	TEXT (N 字节)		
UART 示例	例如, 把数据存储到第 0 块缓存开始的位置: (0xBF 0xC6 0xB4 0xF3 0xD1 0xB6 0xB7 0xC9):							
	0xFD	0x00	0x0A	0x31	0x00	0xBF	0xC6	0xB4
	0xF3	0xD1	0xB6	0xB7	0xC9			

◇ Cplay 命令

名称	发送的数据	说明					
命令字	0x32	把下位机缓存的数据进行 TTS 合成。下位机会从第 0 个缓存块开始进行 TTS 合成					
参数列表	[PAGE] 代码页	4bit, 和 REPEAT 共同组成一个字节, PAGE 占低 4 位。 支持如下代码页: 【0x00】GB2312; 【0x01】GBK; 【0x02】BIG5; 【0x03】UTF16LE 【0x04】UTF8; 【其它】无效值, 下位机将返回错误					
	[REPEAT]重复次数	4bit, 表示重复进行 TTS 合成的次数。和 PAGE 共同组成一个字节, REPEAT 占高 4 位。					
命令帧格式结构	标记	帧数据长度		命令字	参数		
	0xFD	0x00	0xLL	0x32	PAGE_REPEAT		
UART 示例	例如, 用 UTF-8 代码页来对下位机缓存的数据进行 TTS 合成, 重复 2 次。						

	0xFD	0x00	0x02	0x32	0x24
--	------	------	------	------	------

## 2.6 上传协议介绍

芯片通过上传状态帧给上位机反馈状态信息，上传状态帧的格式定义就是上传协议。上传状态帧只包含状态字。状态字由一个字节构成，用来指示通信状态和下位机的运行状态。UART 接口的上传状态帧同下表。

状态字	说明	
0x41	通信状态反馈	表示收到正确的下传命令帧
0x45		表示下传命令帧的命令字错误
0x46		表示下传命令帧的数据不完整，或者帧数据长度错误
0x47		表示下传命令帧的标记错误
0x4A	系统状态反馈	表示下位机系统初始化完成
0x4E		表示下位机正忙
0x4F		表示下位机空闲
其他值	无效值	

## 2.7 上传状态帧

名称	发送的数据	说明
状态字	-	<p>芯片反馈状态给上位机。</p> <p>通信状态反馈：</p> <p>【0x41】表示收到正确的下传命令帧；</p> <p>【0x45】表示下传命令帧的命令字错误；</p> <p>【0x46】表示下传命令帧的数据不完整，或帧数据长度错误</p> <p>【0x47】表示下传命令帧的标记错误；</p> <p>系统状态反馈：</p> <p>【0x4A】表示下位机系统初始化完成；</p> <p>【0x4E】表示下位机正忙；</p> <p>【0x4F】表示下位机空闲；</p> <p>【其它】无效值；</p>
状态帧格式	状态字	
	STA	

UART 示例	例如, 芯片的系统初始化完成后, 反系统反馈 <b>0x4A</b> 给上位机
	<b>0x4A</b>

### 3. 文本标注方法

为了确保 TTS 合成效果, 以下提供一套标注方式, 用户可以通过标注的方式, 达到想要的合成效果, 修复机器合成在自然度、流畅度方面的不足。

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。需要注意的是: 文本标记符全部是半角, 字母必须是小写的英文字母, 不符合要求的不作为文本标记。

设置文本标记时, 标记的位置很重要, 除标记 [n\*]、[r\*]、[o\*]、[y\*] 外, 其余在标记处会强制分句处理, 因此在使用时, 尽量将其放在句首位置。

加标记不会影响整个合成播报内容的变动。如: [w0] 是取消词语间停顿的标记, 合成过程中标记不会被合成。

文本控制标记列表

序号	类型	格式	参数	说明
1	设置标点符号读法	[b*]	*=0/1 0 - 不读 1 - 朗读	设置标点符号是否朗读的标记, 默认为 0 (不读)。
2	设置合成场景	[c*]	*=0/1/2/3/4/5 0-默认场景 1-阅读 (散文类) 2-小说 3-新闻 4-广告 5-交互	
3	恢复默认的合成参数	[d]	---	不能恢复语种和发音人, 其他参数都可以恢复
4	设置音效模式	[e*]	*=0/1/2/3/4/5/6 0 - 关闭 1 - 忽远忽近 2 - 回声 3 - 机器人 4 - 合唱 5 - 水下 6 - 混响	默认为关闭
5	选择语种环境	[g*]	*=0/1/2 0 - 自动判断	默认语种为自动判断。

			1 - 本语种环境 2 - 英语语种环境	
6	设置单词发音方式	[h*]	*=0/1/2 0 - 自动判断单词发音方式 1 - 字母发音方式 2 - 单词发音方式	默认单词为自动判断
7	设置数字处理策略	[n*]	*=0/1/2 0 - 自动判断 1 - 数字作号码处理 2 - 数字作数值处理	默认为自动判断
8	设置汉语号码中“1”的读法	[y*]	*=0/1 0 - 合成号码时“1”读成“yāo”	
9	静音一段时间	[p*]	*=静音的时间长度，无符号整数，如[p500]单位：毫秒(ms)	
10	设置姓名读音策略	[r*]	*=0/1 0 - 自动判断姓氏读音 1 - 强制使用姓氏读音规则	默认为自动判断姓名读音，[r1]标记仅对紧跟其后的姓氏有效，如“[r1]单”自动判断为“shàn”
11	设置音量	[v*]	(*=0~10)	* - 音量值 默认音量为 5，10 对应到最高音量。
12	设置语速	[s*]	*=0~10	默认语速值为 5，最大语速为 10
13	设置语调	[te*]	*=0~100	默认语调值为 50，最大语调为 100
14	强制停顿边界	[w*]	*=0/1/3 0 - 强制标记位置位 L0 边界，无停顿 1 - 强制标记位置位 L1（韵律词）边界，短停顿 3 - 强制标记位置	——

			位 L3 (韵律短语) 边界, 长停顿	
15	设置韵律标注处理策略	[z*]	*=0/1 0 - 不处理韵律标注 1 - 处理韵律标注	默认不处理韵律标注。韵律标注使用“*”标出 L1 划分位置, 使用“#”标出 L3 划分位置。
16	为单个汉字/单词强制指定拼音/音标	[=*]	*=拼音/音标 为前一个汉字/单词设定的拼音/音标	汉字: 声调用后接一位数字 1~5 分别表示阴平、阳平、上声、去声和轻声 5 个声调。 6~9 分别对应阴平、阳平、上声、去声的轻 度, 该标记只能放在非汉语拼音的汉语音节之后指定拼音, 连续出现时以最后一个为准。 示例: “着 [zhuo2]手”, “着”字将读作 “zhuó”

**说明:** 系统默认有自己的合成规则, 会对常见的单位缩写进行单位拼读, 如 “5.5t” 合成 “5.5 吨”, 如需合成另外的读法, 请以中文方式或参考第 15 条音标/拼音注释方式合成。