

语音合成播报模块通讯协议

一、通讯方式

模块通讯方式为 I2C 通讯，模块 I2C 地址为 0x30（旧版为 0x50），下述数据帧发送均通过 I2C 通讯实现。

二、语音播报

1. 语音播报功能支持的控制命令

上位机发送给语音识别模块的所有语音播报命令都需要用“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

下列为命令帧说明，帧头尾固定值 0xFD，数据长度分为两个字节发送，最后为数据区，大小小于 4k。

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

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播报芯片的各种功能。不同命令字有对应不同的命令参数。具体见后续的说明。

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

2. 语音播报命令

语音播报通过 I2C 将帧头固定值为 0xFD，数据区长度分为两个字节，数据区长度为带合成文本长度+2（命令字和文本格式各为一个字节），需要注意文本长度表示的是文本的字节长度，例如文本编码为 GB2312，字符串“123”的长度为 3，字符串“你好”的长度为 4，因为 GB2312 中文一个汉字为 2 个字节的长度。普通的英文和数字均为 1 个字节，具体字符的编码可以到网上看具体数值和长度。

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

特别提示:当语音合成播报模块正在合成文本的时候,如果又收到一帧有效合成命令帧,芯片会立刻停止当前正在合成的文本,转而合成新收到的文本,这里就需要利用查询芯片状态的命令,只有当检测到芯片空闲的时候再发送新的命令帧。

3. 停止合成命令

通过 I2C 将以下四个数据发送给模块可以使模块停止合成。发送数据可以在一个起始和停止之间一口气发送完 4 个数据,但是为了避免连续发送数据过快可能造成内部数据处理不够及时导致出错,建议一个起始位和一个停止位之间发送一个数据,并适当留一点点间隔时间。

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		
	高字节	低字节	命令字		
	0x00	0x01	0x02		

4. 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		
	高字节	低字节	命令字		
	0x00	0x01	0x03		

4. 恢复合成

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

5. 芯片状态读取

通过 I2C 通讯将以下 4 个数据发送给模块，之后再通过 I2C 通讯读取一个字节，获取芯片状态值。发送数据可以在一个起始和停止之间一口气发送完 4 个数据，但是为了避免连续发送数据过快可能造成内部数据处理不够及时导致出错，建议一个起始位和一个停止位之间发送一个数据，并适当留一点点间隔时间。读取的时候需要一个完整的 I2C 时序，即 S, I2CADDR+1 (读), ACK, DATA, NACK, P。默认状态值为 0xff。

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
	高字节	低字节		命令字
	0xFD	0x00	0x01	0x21

除了默认值 0xff 以外，其他芯片状态值为下表，在出现以下情况后内部协处理器会将芯片状态值更新为对应状态，在接收到读取后发送给上位机，之后重新更新为默认值 0xff，等待芯片回传新的状态值再更新对应的状态。

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

三、文本标记控制

1. 文本控制标记的使用方法

语音播报功能支持多种文本控制标记，可以满足用户对语音播报发音人、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。

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”，即根据上述的语音播报中的待合成文本编写为“[v3]”则设置音量 3 级，编写为“[v3]我在小声说话，[v10]我在大声说话”，则设置音量同时合成对应的语音进行播报。

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”。

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	[m3]
		? 为 3，设置发音人为小燕(女声，推荐发音人)	
		? 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	[n0]
		? 为 1，数字作号码处理	
		? 为 2，数字作数值处理	
数字“0”在读作英文、号码时的读法	[o?]	? 为 0，读成“zero”	[o0]
		? 为 1，读成“欧”音	
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	[r0]
		? 为 1，强制使用姓氏读音规则	
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]
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	
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”	
是否使用韵律标记“*”和“#”	[z?]	? 为 0，“*”和“#”读出符号	[z0]
		? 为 1，处理成韵律，“*”用于断词，“#”用于停顿	
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	

使用文本控制标记需注意：

- 1、所有的控制标识均为半角字符。
- 2、控制标识需要按照语音播报命令的格式发送，控制标识作为文本进行合成，即合成命令是“帧头 + 数据区长度 + 合成命令字 + 文本编码格式 + 控制标识文本”的格式。
- 3、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。

4、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。

3. 控制标记使用示例

3.1[f?]标记的使用

“[f0]亚博智能”合成为“亚-博-智-能”，中间会出现停顿。“[f1]亚博智能”合成为“亚博智能”。

3.2[g?]标记的使用

“[g0]120km”合成为“一百二十公里”，自动判断合成为中文播报，“[g1]120km”合成为：“一百二十公里”，指定合成为中文。“[h0][g2]120km”合成为“One hundred and twenty kilometers”，指定合成英文并指定英文按单词和播报。

3.3[h?]标记的使用

“[h0]你的英语是 you”合成为“你的英语是 you”，自动判断单词发音，“[h1]你的英语是 you”合成为：“你的英语是 y-o-u”，指定字母发音。“[h2]你的英语是 you”合成为“你的英语是 you”，指定单词发音。

3.4[i?]标记的使用

“[i0]亚博智 neng2”合成为“亚博智 NENG”，不识别拼音格式，“[i1]亚博智 neng”合成为“亚博智能”，识别拼音格式。

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。

3.5[m?]标记的使用

“[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。”唐老鸭模拟器合成：“我是唐老鸭”，“[m55]我是许小宝。”女童声合成：“我是许小宝”。

3.6[n?]标记的使用

“[n0]234343545。”模块自动判断，读作：二三三四三四五五五，“[n1]234343545。”模块强制按照号码的方式合成数字串，读作：二三三四三四五五五，“[n2]234343545。”模块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。

3.7[o?]标记的使用

“[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。

3.8[p?]标记的使用

“欢迎使用用[p500]亚博智能飞[p1000]语音合成播报模块”模块在“欢迎使用”合成完毕后静音 500 毫秒，在“智能”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。

3.9[r?]标记的使用

“[r0] 查丽”模块按照默认模块按照默认的方式合成。读作：chá lì。“[r1] 查丽”模块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。

3.10[s?]标记的使用

“您好，[s9]欢迎使用亚博智能[s5] 语音合成播报模块”模块语速的调节功能，提供 11 级的语速调节，模块默认的语速为 5。本句合成时“您好”为 5 级语速，“欢迎使用亚博智能”为 9 级语速，“语音合成播报模块”为 5 级语速。

3.11[t?]标记的使用

“您好，[t10]欢迎使用亚博智能[s0] 语音合成播报模块”模块语调的调节功能，提供 11 级的语调调节，模块默认的语调大小为 5，最小为 0，最大为 10。本句合成时“您好”为 5 级语调，“欢迎使用亚博智能”为 10 级语调，“语音合成播报模块”为 0 级语调。

3.12[v?]标记的使用

“欢迎[v3]使用[v8]亚博智能语音合成播报模块”模块音量的调节功能，提供 11 级的音量调节，最小 0 为静音，模块默认的音量大小为 5。本句合成时“欢迎”为 5 级音量，“使用”为 3 级音量，“亚博智能语音合成播报模块”为 8 级音量。

3.13[x?]标记的使用

“提示音铃声[x0]sound101 对应的声音是[x1]sound101”合成为“提示音铃声[x0]sound101 对应的声音是”然后播放提示音，设置 1 则后续的 sound101 以内置的对应的提示音进行播报，设置 0 则按原来的英文数字混读播报。用户只需在平常的文本前面加上[x1]和对应的内置提示音对应字符串即可播放对应的提示音。提示音列表见下图。

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

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

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
----------	----------	----------	----------	----------

3.14[y?]标记的使用

“[y0]010-58301005-8016”模块按照“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。“[y0]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。

注意：此标记必须是在合成号码类型文本的时候才有效。

3.15[z?]标记的使用

“[z0]深圳#亚博*智能”模块读成“深圳井号亚博星号智能”，“[z1]深圳#亚博*智能”，芯片读成“深圳 亚博 智能”，#和*的地方被处理成短暂的停顿。

3.16[=?]标记的使用

“曾[=deng1]国藩”模块强制的将“曾”按照“dēng”音合成播报。

3.17[d]标记的使用

“[s2][v3][t10]亚博智能，[d]亚博智能”前一个“亚博智能”按照 2 级语速、3 级音量、10 级语调合成；后一个“亚博智能”恢复到默认的语速、音量、语调进行合成。