

零死角玩转STM32



LTDC—液晶显示中 英文

淘宝：firestm32.taobao.com

论坛：www.chuxue123.com



扫描进入淘宝店铺

主讲内容



01

字符编码

02

什么是字模？

03

各种模式的液晶显示字符实验

参考资料:《零死角玩转STM32》

“LTDC—液晶显示中英文” 章节

LTDC—液晶显示中英文



本章字符编码说明参考网站

字符编码及转换测试: <http://www.qqxiuzi.cn/daohang.htm>

Unicode官网: <http://www.unicode.org>。

字符编码

由于计算机只能识别0和1，文字也只能以0和1的形式在计算机里存储，所以我们需要对文字进行编码才能让计算机处理，编码的过程就是规定特定的01数字串来表示特定的文字，最简单的字符编码例子是ASCII码。

ASCII编码

在程序设计中使用ASCII编码表约定了一些控制字符、英文及数字。它们在存储器中，本质也是二进制数，只是我们约定这些二进制数可以表示某些特殊意义，如以ASCII编码解释数字“0x41”时，它表示英文字符“A”。

LTDC—液晶显示中英文



ASCII编码表

十进制	十六进制	缩写/字符	解释
0	0	NUL(null)	空字符
1	1	SOH(start of headline)	标题开始
2	2	STX (start of text)	正文开始
3	3	ETX (end of text)	正文结束
4	4	EOT (end of transmission)	传输结束
5	5	ENQ (enquiry)	请求
6	6	ACK (acknowledge)	收到通知
7	7	BEL (bell)	响铃
8	8	BS (backspace)	退格
9	9	HT (horizontal tab)	水平制表符
10	0A	LF (NL line feed, new line)	换行键
11	0B	VT (vertical tab)	垂直制表符
12	0C	FF (NP form feed, new page)	换页键
13	0D	CR (carriage return)	回车键
14	0E	SO (shift out)	不用切换
15	0F	SI (shift in)	启用切换
16	10	DLE (data link escape)	数据链路转义
17	11	DC1 (device control 1)	设备控制1
18	12	DC2 (device control 2)	设备控制2
19	13	DC3 (device control 3)	设备控制3
20	14	DC4 (device control 4)	设备控制4
21	15	NAK (negative acknowledge)	拒绝接收
22	16	SYN (synchronous idle)	同步空闲
23	17	ETB (end of trans. block)	传输块结束
24	18	CAN (cancel)	取消
25	19	EM (end of medium)	介质中断
26	1A	SUB (substitute)	替补
27	1B	ESC (escape)	换码(溢出)
28	1C	FS (file separator)	文件分割符
29	1D	GS (group separator)	分组符
30	1E	RS (record separator)	记录分离符
31	1F	US (unit separator)	单元分隔符

ASCII编码表

十进制	十六进制	缩写/字符		十进制	十六进制	缩写/字符
32	20	(space)空格		80	50	P
33	21	!		81	51	Q
34	22	"		82	52	R
35	23	#		83	53	S
36	24	\$		84	54	T
37	25	%		85	55	U
38	26	&		86	56	V
39	27	'		87	57	W
40	28	(88	58	X
41	29)		89	59	Y
42	2A	*		90	5A	Z
43	2B	+		91	5B	[
44	2C	,		92	5C	\
45	2D	-		93	5D]
46	2E	.		94	5E	^
47	2F	/		95	5F	_
48	30	0		96	60	`
49	31	1		97	61	a
50	32	2		98	62	b
51	33	3		99	63	c
52	34	4		100	64	d
53	35	5		101	65	e

54	36	6		102	66	f
55	37	7		103	67	g
56	38	8		104	68	h
57	39	9		105	69	i
58	3A	:		106	6A	j
59	3B	;		107	6B	k
60	3C	<		108	6C	l
61	3D	=		109	6D	m
62	3E	>		110	6E	n
63	3F	?		111	6F	o
64	40	@		112	70	p
65	41	A		113	71	q
66	42	B		114	72	r
67	43	C		115	73	s
68	44	D		116	74	t
69	45	E		117	75	u
70	46	F		118	76	v
71	47	G		119	77	w
72	48	H		120	78	x
73	49	I		121	79	y
74	4A	J		122	7A	z
75	4B	K		123	7B	{
76	4C	L		124	7C	
77	4D	M		125	7D	}
78	4E	N		126	7E	~
79	4F	O		127	7F	DEL (delete) 删除

LTDC—液晶显示中英文



ASCII编码

ASCII码表分为两部分，第一部分是控制字符或通讯专用字符，它们的数字编码从0~31，它们并没有特定的图形显示，但会根据不同的应用程序，而对文本显示有不同的影响。ASCII码的第二部分包括空格、阿拉伯数字、标点符号、大小写英文字母以及“DEL（删除控制）”，这部分符号的数字编码从32~127，除最后一个DEL符号外，都能以图形的方式来表示，它们属于传统文字书写系统的一部分。

后来，计算机引进到其它国家的时候，由于他们使用的不是英语，他们使用的字母在ASCII码表中没有定义，所以他们采用127号之后的位来表示这些新的字母，还加入了各种形状，一直编号到255。从128到255这些字符被称为ASCII扩展字符集。至此基本存储单位Byte(char)能表示的编号都被用完了。

LTDC—液晶显示中英文



中文编码

英文书写系统都是由**26**个基本字母组成，利用**26**个字母组可合出不同的单词，所以用**ASCII**码表就能表达整个英文书写系统。而中文书写系统中的汉字是独立的方块，若参考单词拆解成字母的表示方式，汉字可以拆解成部首、笔画来表示，但这样会非常复杂(可参考五笔输入法编码)，所以中文编码直接对方块字进行编码，一个汉字使用一个号码。

由于汉字非常多，常用字就有**6000**多个，如果像**ASCII**编码表那样只使用**1**个字节最多只能表示**256**个汉字，所以我们使用**2**个字节来编码。

LTDC—液晶显示中英文



1. GB2312标准

我国首先定义的是GB2312标准。它把ASCII码表127号之后的扩展字符集直接取消掉，并规定小于127的编码按原来ASCII标准解释字符。当2个大于127的字符连在一起时，就表示1个汉字，第1个字节使用 (0xA1-0xFE) 编码，第2个字节使用(0xA1-0xFE)编码，这样的编码组合起来可以表示了7000多个符号，其中包含6763个汉字。在这些编码里，我们还把数学符号、罗马字母、日文假名等都编进表中，就连原来在ASCII里原本就有的数字、标点以及字母也重新编了2个字节长的编码，这就是平时在输入法里可切换的“全角”字符，而标准的ASCII码表中127号以下的就被称为“半角”字符。

LTDC—液晶显示中英文



1. GB2312标准

下表说明了GB2312是如何兼容ASCII码的，当我们设定系统使用GB2312标准的时候，它遇到一个字符串时，会按字节检测字符值的大小，若遇到连续两个字节的数值都大于127时就把这两个连续的字节合在一起，用GB2312解码，若遇到的数值小于127，就直接用ASCII把它解码。

第1字节	第2字节	表示的字符	说明
0x68	0x69	(hi)	两个字节的值都小于127(0x7F)，使用ASCII解码
0xB0	0xA1	(啊)	两个字节的值都大于127(0x7F)，使用GB2312解码

17区

17	0	1	2	3	4	5	6	7	8	9
0	薄	雹	保	堡	饱	宝	抱	报	暴	
1	豹	鲍	杯	碑	悲	卑	北	鞑	背	
2	贝	钹	狐	备	惫	焙	被	奔	苯	
3	本	笨	绷	胥	泵	蹦	迸	逼	鼻	
4	比	鄙	彼	碧	蓖	蔽	毕	毙	臂	
5	币	庇	闭	敝	弊	必	辟	壁	慝	
6	避	陛	鞭	编	贬	扁	便	变	慝	
7	辨	辩	辨	编	彪	膘	表	鳖	冰	
8	别	瘰	彬	濒	濒					
9	柄	丙	秉	饼	炳					

LTDC—液晶显示中英文



区位码

在GB2312编码的实际使用中，有时会用到区位码的概念。GB2312编码对所收录字符进行了“分区”处理，共94个区，每区含有94个位，共8836个码位。而区位码实际是GB2312编码的内部形式，它规定对收录的每个字符采用两个字节表示，第一个字节为“高字节”，对应94个区；第二个字节为“低字节”，对应94个位。所以它的区位码范围是：0101-9494。为兼容ASCII码，区号和位号分别加上0xA0偏移就得到GB2312编码。在区位码上加上0xA0偏移，可求得GB2312编码范围：0xA1A1—0xFEFE，其中汉字的编码范围为0xB0A1-0xF7FE，第一字节0xB0-0xF7（对应区号：16—87），第二个字节0xA1-0xFE（对应位号：01—94）。

例如，“啊”字是GB2312编码中的第一个汉字，它位于16区的01位，所以它的区位码就是1601，加上0xA0偏移，其GB2312编码为0xB0A1。其中区位码为0101的码位表示的是“空格”符。

2. GBK编码

据统计，GB2312编码中表示的6763个汉字已经覆盖中国大陆99.75%的使用率，单看这个数字已经很令人满意了，但是不能因为那些文字不常用就不让它进入信息时代，而且生僻字在人名、文言文中的出现频率是非常高的。

为此我们在GB2312标准的基础上又增加了14240个新汉字(包括所有后面介绍的Big5中的所有汉字)和符号，这个方案被称为GBK标准。增加这么多字符，按照GB2312原来的格式来编码，2个字节已经存储不下，我们的程序员修改了一下格式，不再要求第2个字节的编码值必须大于127，只要第1个字节大于127就表示这是一个汉字的开始，这样就做到兼容ASCII和GB2312标准了。

LTDC—液晶显示中英文



2. GBK编码

说明了GBK是如何兼容ASCII和GB2312标准的，当我们设定系统使用GBK标准的时候，它按顺序遍历字符串，按字节检测字符值的大小，若遇到一个字符的值大于127时，就再读取它后面的一个字符，把这两个字符值合在一起，用GBK解码，解码完后，再读取第3个字符，重新开始以上过程，若该字符值小于127，则直接用ASCII解码。

第1字节	第2字节	第3字节	表示的字符	说明
0x68(<7F)	0xB0(>7F)	0xA1(>7F)	(h啊)	第1个字节小于127，使用ASCII解码，第2个字节大于127，直接使用GBK解码，兼容GB2312
0xB0(>7F)	0xA1(>7F)	0x68(<7F)	(啊h)	第1个字节大于127，直接使用GBK码解释，第3个字节小于127，使用ASCII解码
0xB0(>7F)	0x56(<7F)	0x68(<7F)	(癡h)	第1个字节大于127，第2个字节虽然小于127，直接使用GBK解码，第3个字节小于127，使用ASCII解码

LTDC—液晶显示中英文



2. GBK编码

说明了GBK是如何兼容ASCII和GB2312标准的，当我们设定系统使用GBK标准的时候，它按顺序遍历字符串，按字节检测字符值的大小，若遇到一个字符的值大于127时，就再读取它后面的一个字符，把这两个字符值合在一起，用GBK解码，解码完后，再读取第3个字符，重新开始以上过程，若该字符值小于127，则直接用ASCII解码。

第1字节	第2字节	第3字节	表示的字符	说明
0x68(<7F)	0xB0(>7F)	0xA1(>7F)	(h啊)	第1个字节小于127，使用ASCII解码，第2个字节大于127，直接使用GBK解码，兼容GB2312
0xB0(>7F)	0xA1(>7F)	0x68(<7F)	(啊h)	第1个字节大于127，直接使用GBK码解释，第3个字节小于127，使用ASCII解码
0xB0(>7F)	0x56(<7F)	0x68(<7F)	(癡h)	第1个字节大于127，第2个字节虽然小于127，直接使用GBK解码，第3个字节小于127，使用ASCII解码

LTDC—液晶显示中英文



3. GB18030

随着计算机技术的普及，我们后来又在**GBK**的标准上不断扩展字符，这些标准被称为**GB18030**，如**GB18030-2000**、**GB18030-2005**等(“-”号后面的数字是制定标准时的年号)，**GB18030**的编码使用4个字节，它利用前面标准中的第2个字节未使用的“0x30-0x39”编码表示扩充四字节的后缀，兼容**GBK**、**GB2312**及**ASCII**标准。**GB18030-2000**主要在**GBK**基础上增加了“**CJK(中日韩)统一汉字扩充A**”的汉字。加上前面**GBK**的内容，**GB18030-2000**一共规定了**27533**个汉字（包括部首、部件等）的编码，还有一些常用非汉字符号。

GB18030-2005的主要特点是在**GB18030-2000**基础上增加了“**CJK(中日韩)统一汉字扩充B**”的汉字。增加了**42711**个汉字和多种我国少数民族文字的编码（如藏、蒙古、傣、彝、朝鲜、维吾尔文等）。加上前面**GB18030-2000**的内容，一共收录了**70244**个汉字。

LTDC—液晶显示中英文



各个标准的对比说明

GB2312、GBK及GB18030是汉字的国家标准编码，新版向下兼容旧版，各个标准简要说明见下表，目前比较流行的是**GBK**编码，因为每个汉字只占用**2**个字节，而且它编码的字符已经能满足大部分的需求，但国家要求一些产品必须支持**GB18030**标准。



类别	编码范围	汉字编码范围	扩充汉字数	说明
GB2312	第一字节 0xA1-0xFE 第二字节 0xA1-0xFE	第一字节 0xB0-0xF7 第二字节 0xA1-0xFE	6763	除汉字外，还包括拉丁字母、希腊字母、日文平假名及片假名字母、俄语西里尔字母在内的 682 个全角字符
GBK	第一字节 0x81-0xFE 第二字节 0x40-0xFE	第一字节 0x81-0xA0 第二字节 0x40-0xFE	6080	包括部首和构件，中日韩汉字，包含了 BIG5 编码中的所有汉字，加上 GB2312 的原内容，一共有 21003 个汉字
		第一字节 0xAA-0xFE 第二字节 0x40-0xA0	8160	
GB18030-2000	第一字节 0x81-0xFE 第二字节 0x30-0x39 第三字节 0x81-0xFE 第四字节 0x30-0x39	第一字节 0x81-0x82 第二字节 0x30-0x39 第三字节 0x81-0xFE 第四字节 0x30-0x39	6530	在 GBK 基础上增加了中日韩统一汉字扩充 A 的汉字，加上 GB2312、GBK 的内容，一共有 27533 个汉字
GB18030-2005	第一字节 0x81-0xFE 第二字节 0x30-0x39 第三字节 0x81-0xFE	第一字节 0x95-0x98 第二字节 0x30-0x39 第三字节 0x81-0xFE	42711	在 GB18030-2000 的基础上增加了 42711 中日韩统一
	第四字节 0x30-0x39	第四字节 0x30-0x39		汉字扩充 B 中的汉字和多种我国少数民族文字的编码（如藏、蒙古、傣、彝、朝鲜、维吾尔文等），加上前面 GB2312、GBK、GB18030-2000 的内容，一共 70244 个汉字

LTDC—液晶显示中英文



4. Big5编码

在台湾、香港等地区，使用较多的是Big5编码，它的主要特点是收录了繁体字。而从GBK编码开始，已经把Big5中的所有汉字收录进编码了。即对于汉字部分，GBK是Big5的超集，Big5能表示的汉字，在GBK都能找到那些字相应的编码，但他们的编码是不一样的，两个标准不兼容，如GBK中的“啊”字编码是“0xB0A1”，而Big5标准中的编码为“0xB0DA”。

零死角玩转STM32



THANKS

论坛：www.chuxue123.com

淘宝：firestm32.taobao.com



扫描进入淘宝店铺