

GT20L16P1Y 标准汉字字库芯片

用户手册 DATASHEET

- n 支持 GB2312 简体汉字字符集
- n 支持 UNICODE 常用简繁体汉字字符集
- n 支持 Unicode 175 国中外文字符集
拉丁文、希腊文、西里尔文、阿拉伯文、希伯来文、泰文 等
- n 排置方式：竖置横排
- n 总线接口：SPI 串行总线
- n 芯片形式：SOT23-6 封装

VER 1.0

2010-Q4

目 录

第一部分：硬件部分

1 概述	3
1.1 芯片特点.....	3
1.2 芯片内容.....	4
1.3 字型样张.....	5
2 引脚描述与接口连接	6
2.1 引脚描述.....	3
2.2 HOST CPU 主机 SPI 接口电路示意图.....	6
3 操作指令	7
3.1 指令参数.....	7
3.2 Read Data Bytes（一般读取）.....	7
3.3 Read Data Bytes at Higher Speed（快速读取点阵数据）.....	8
4 电气特性	9
4.1 绝对最大额定值.....	9
4.2 DC 特性.....	9
4.3 AC 特性.....	9
5 封装尺寸	11

第二部分：软件部分

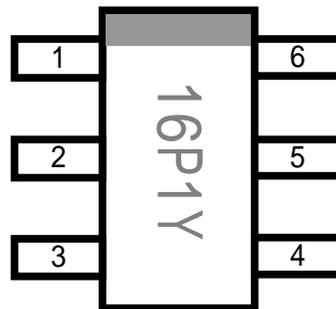
6 字库调用方法	12
6.1 字符点阵排列格式.....	12
6.2 点阵字库地址表.....	14
6.3 字符在芯片中的地址计算方法.....	14
7 附录	18
7.1 Unicode 字符区字符.....	18
7.2 175 国中外文字库拼音索引表.....	25
7.3 175 国中外文字库字母索引表.....	27
7.4 175 国中外文字库总表.....	29
7.5 UTF8 编码到 UTF16 编码转换程序.....	33

1 概述

GT20L16P1Y是一款15X16汉字点阵和16点外文点阵字库芯片，支持GB2312字符集、UNICODE字符集,同时支持拉丁文、希腊、基里尔文、阿拉伯文、希伯来文、泰文。排列格式为竖置横排。用户通过字符内码，利用本手册提供的方法计算出该字符点阵在芯片中的地址，可从该地址连续读出字符点阵信息。

1.1 芯片特点

- 数据总线：SPI 串行总线接口
- 点阵排列方式：字节竖置横排
- 时钟频率：30MHz(max.) @3.3V
- 工作电压：2.2V~3.6V
- 电流：
 - 工作电流：8mA
 - 待机电流：8uA
- 封装：SOT23-6
- 尺寸 SOT23-6：2.9mmX1.6 mm x1.10mm
- 工作温度：-20℃~85℃



1.2 引脚描述

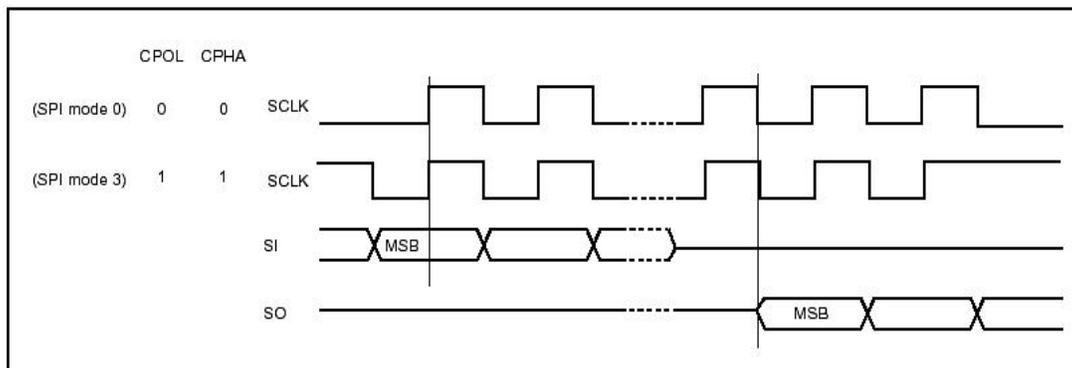
SOT23-6	名称	I/O	描述
1	SCLK	I	串行时钟输入 (Serial clock input)
2	GND		地(Ground)
3	CS#	I	片选输入 (Chip enable input)
4	VCC		电源(+ 3.3V Power Supply)
5	SO	O	串行数据输出 (Serial data output)
6	SI	I	串行数据输入 (Serial data input)

串行数据输出 (SO)：该信号用来把数据从芯片串行输出，数据在时钟的下降沿移出。

串行数据输入 (SI)：该信号用来把数据从串行输入芯片，数据在时钟的上升沿移入。

串行时钟输入 (SCLK)：数据在时钟上升沿移入，在下降沿移出。

片选输入 (CS#)：所有串行数据传输开始于CE#下降沿，CE#在传输期间必须保持为低电平，在两条指令之间保持为高电平。



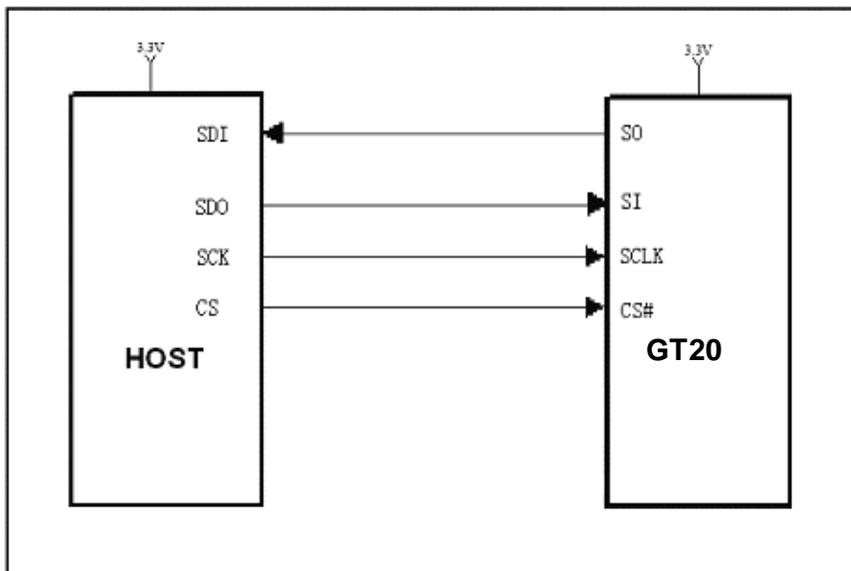
1.3 芯片内容

字符集		点阵 字符数	等宽字符		不等宽字符
			8X16	15X16	16 点
ASCII 字符			96		
GB2312/Unicode	汉字			5271+3417	
Unicode	拉丁文系 (130 国)	Basic(基本)	96		
		Supplement (补充)	96		
		Extended A (扩展 A)	128		
		Extended B (扩展 B)	80		
		Extended Additional (扩展附加)	96		
	希腊文系 (2 国)	Basic(基本)	96		
	西里尔文系 (15 国)	Basic(基本)	208		
	阿拉伯文系 (24 国)	Basic(基本)			256
		Form A(附表 A)			176
		Form B(附表 B)			144
	希伯来文 (1 国)	Basic(基本)	112		
	泰文 (1 国)	Basic(基本)	128		

2 引脚描述与接口连接

2.1 HOST CPU 主机 SPI 接口电路示意图

SPI 与主机接口电路连接可以参考下图。



HOST CPU 主机 SPI 接口电路示意图

3 操作指令

3.1 指令参数

Instruction Set

Instruction	Description	Instruction Code(One-Byte)		Address Bytes	Dummy Bytes	Data Bytes
READ	Read Data Bytes	0000 0011	03 h	3	—	1 to ∞
FAST_READ	Read Data Bytes at Higher Speed	0000 1011	0B h	3	1	1 to ∞

所有对本芯片的操作只有 2 个，那就是 Read Data Bytes (READ “一般读取”)和 Read Data Bytes at Higher Speed (FAST_READ “快速读取点阵数据”)。

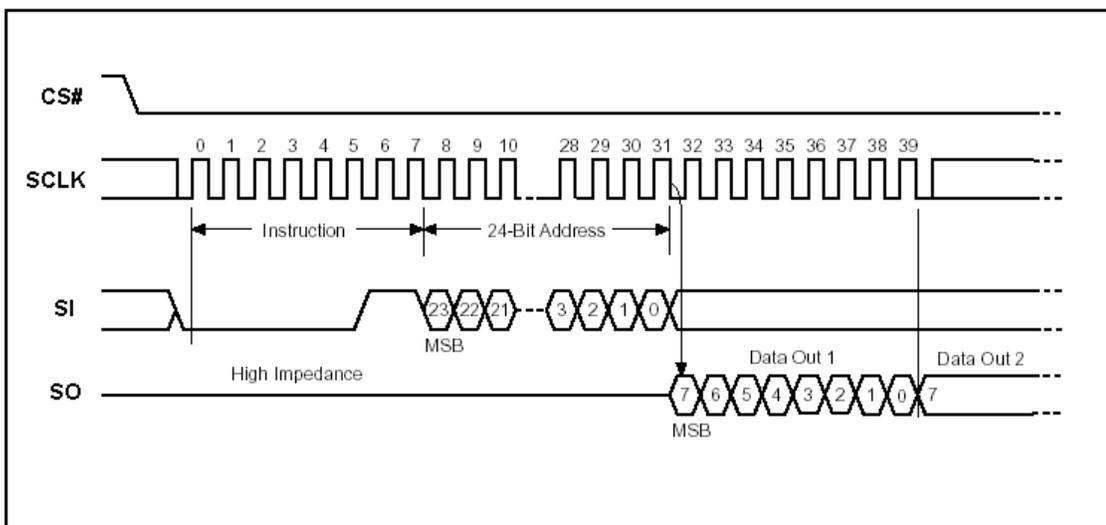
3.2 Read Data Bytes (一般读取)

Read Data Bytes 需要用指令码来执行每一次操作。READ 指令的时序如下(图):

- n 首先把片选信号 (CS#) 变为低，紧跟着的是 1 个字节的命令字 (03 h) 和 3 个字节的地址和通过串行数据输入引脚 (SI) 移位输入，每一位在串行时钟 (SCLK) 上升沿被锁存。
- n 然后该地址的字节数据通过串行数据输出引脚 (SO) 移位输出，每一位在串行时钟 (SCLK) 下降沿被移出。
- n 读取字节数据后，则把片选信号 (CS#) 变为高，结束本次操作。

如果片选信号 (CS#) 继续保持为底，则下一个地址的字节数据继续通过串行数据输出引脚 (SO) 移位输出。

图：Read Data Bytes (READ) Instruction Sequence and Data-out sequence:



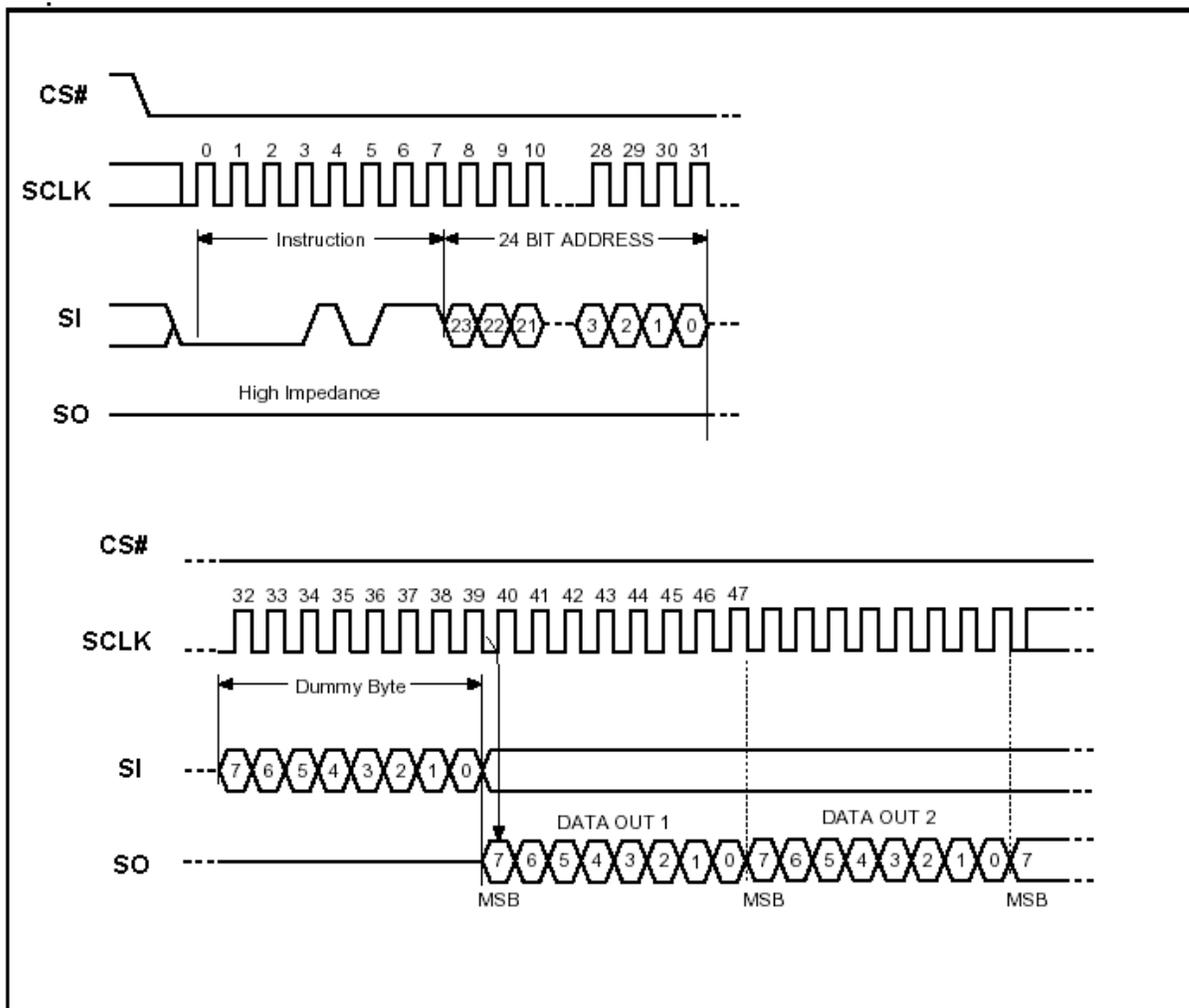
3.3 Read Data Bytes at Higher Speed (快速读取点阵数据)

Read Data Bytes at Higher Speed 需要用指令码来执行操作。READ_FAST 指令的时序如下(图):

- n 首先把片选信号 (CS#) 变为低, 紧跟着的是 1 个字节的命令字 (0B h) 和 3 个字节的地址以及一个字节 Dummy Byte 通过串行数据输入引脚 (SI) 移位输入, 每一位在串行时钟 (SCLK) 上升沿被锁存。
- n 然后该地址的字节数据通过串行数据输出引脚 (SO) 移位输出, 每一位在串行时钟 (SCLK) 下降沿被移出。
- n 如果片选信号 (CS#) 继续保持为底, 则下一个地址的字节数据继续通过串行数据输出引脚 (SO) 移位输出。例: 读取一个 15x16 点阵汉字需要 32Byte, 则连续 32 个字节读取后结束一个汉字的点阵数据读取操作。

如果不需要继续读取数据, 则把片选信号 (CS#) 变为高, 结束本次操作。

图: Read Data Bytes at Higher Speed (READ_FAST) Instruction Sequence and Data-out sequence:



4 电气特性

4.1 绝对最大额定值

Symbol	Parameter	Min.	Max.	Unit	Condition
T _{OP}	Operating Temperature	-20	85	°C	
T _{STG}	Storage Temperature	-65	150	°C	
V _{CC}	Supply Voltage	-0.3	3.6	V	
V _{IN}	Input Voltage	-0.3	V _{CC} +0.3	V	
GND	Power Ground	-0.3	0.3	V	

4.2 DC 特性

Condition: T_{OP} = -20°C to 85°C, GND=0V

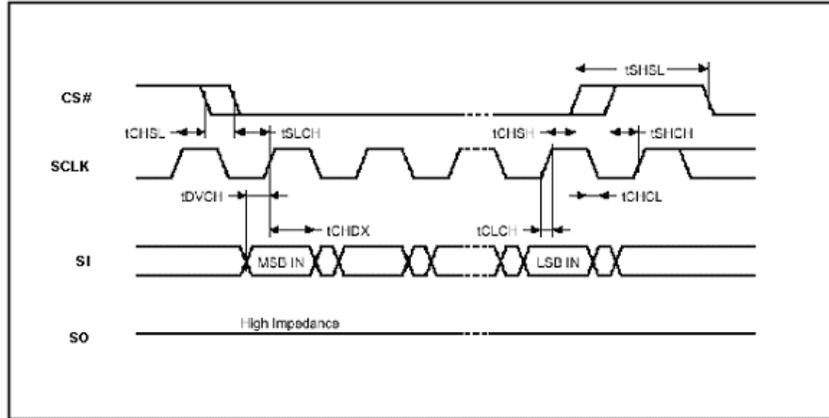
Symbol	Parameter	Min.	Max.	Unit	Condition
I _{DD}	V _{CC} Supply Current(active)		8	mA	V _{CC} =2.2~3.6V
I _{SB}	V _{CC} Standby Current		8	uA	
V _{IL}	Input LOW Voltage	-0.3	0.3V _{CC}	V	
V _{IH}	Input HIGH Voltage	0.7V _{CC}	V _{CC} +0.4	V	
V _{OL}	Output LOW Voltage		0.4 (I _{OL} =1.6mA)	V	
V _{OH}	Output HIGH Voltage	0.8V _{CC} (I _{OH} =-100uA)		V	
I _{LI}	Input Leakage Current	0	2	uA	
I _{LO}	Output Leakage Current	0	2	uA	

Note: I_{IL}: Input LOW Current, I_{IH}: Input HIGH Current,
I_{OL}: Output LOW Current, I_{OH}: Output HIGH Current,

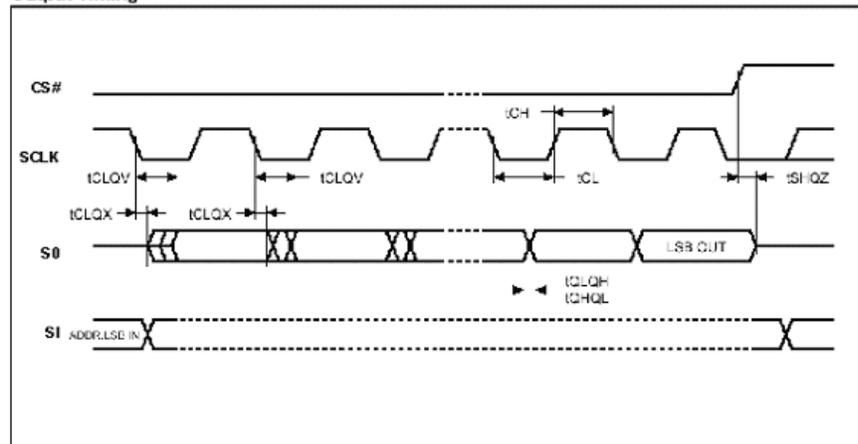
4.3 AC 特性

Symbol	Alt.	Parameter	Min.	Max.	Unit
F _c	F _c	Clock Frequency	D.C.	30	MHz
t _{CH}	t _{CLH}	Clock High Time	15		ns
t _{CL}	t _{CLL}	Clock Low Time	15		ns
t _{CLCH}		Clock Rise Time(peak to peak)	0.1		V/ns
t _{CHCL}		Clock Fall Time (peak to peak)	0.1		V/ns
t _{SLCH}	t _{CSS}	CS# Active Setup Time (relative to SCLK)	5		ns
t _{CHSL}		CS# Not Active Hold Time (relative to SCLK)	5		ns
t _{DVCH}	t _{DSU}	Data In Setup Time	2		ns
t _{CHDX}	t _{DH}	Data In Hold Time	5		ns
t _{CHSH}		CS# Active Hold Time (relative to SCLK)	5		ns
t _{SHCH}		CS# Not Active Setup Time (relative to SCLK)	5		ns
t _{SHSL}	t _{CSH}	CS# Deselect Time	100		ns
t _{SHQZ}	t _{DIS}	Output Disable Time		9	ns
t _{CLQV}	t _V	Clock Low to Output Valid		9	ns
t _{CLQX}	t _{HO}	Output Hold Time	0		ns

Serial Input Timing

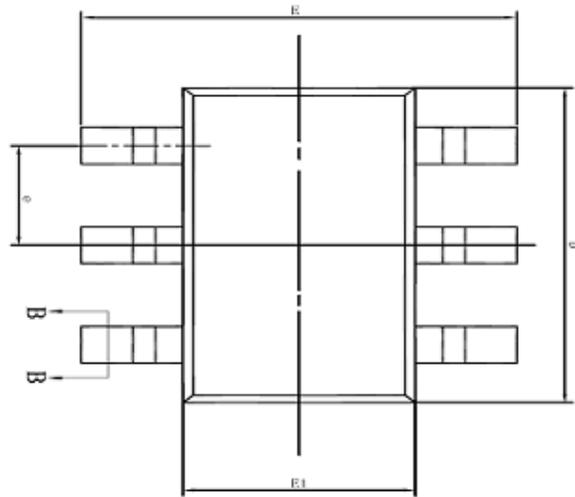
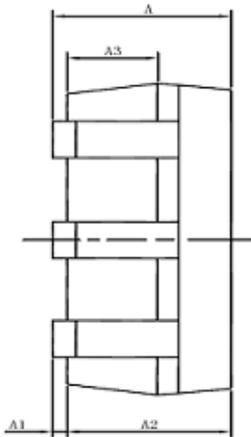


Output Timing

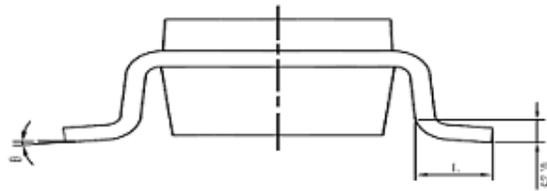


5 封装尺寸

SOT23-6 Package



SYMBOL	MILLIMETER		
	MIN	NOM	MAX
A	—	—	1.30
A1	0.04	0.07	0.10
A2	1.00	1.10	1.20
A3	0.55	0.65	0.75
D	2.72	2.92	3.12
E	2.60	2.80	3.00
E1	1.40	1.60	1.80
e	0.95BSC		
L	0.30	—	0.60
θ	0	—	8°



SOT23-6 封装

6 字库调用方法

6.1 字符点阵排列格式

每个字符在芯片中是以汉字点阵字模的形式存储的，每个点用一个二进制位表示，存 1 的点，当显示时可以在屏幕上显示亮点，存 0 的点，则在屏幕上不显示。点阵排列格式为竖置横排：即一个字节的低位表示下面的点，高位表示上面的点（如果用户按 16bit 总线宽度读取点阵数据，请注意高低字节的顺序），排满一行后再排下一行。这样把点阵信息用来直接在显示器上按上述规则显示，则将出现对应的字符。

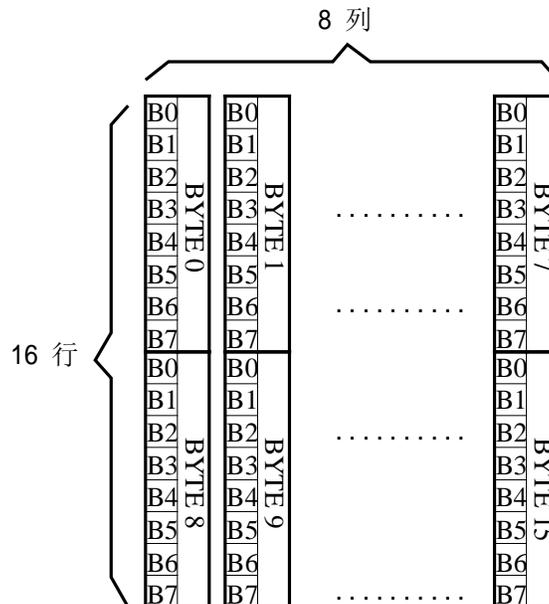
6.1.1 8X16 点字符排列格式

适用于此种排列格式的字体有：

8X16 点 ASCII 字符

8X16 点拉丁文、希腊文、基里尔文字符、希伯来文系字符、泰文字符

8X16 点字符信息需要 16 个字节（BYTE 0 – BYTE15）来表示。该点阵数据是竖置横排的，其具体排列结构如下图：

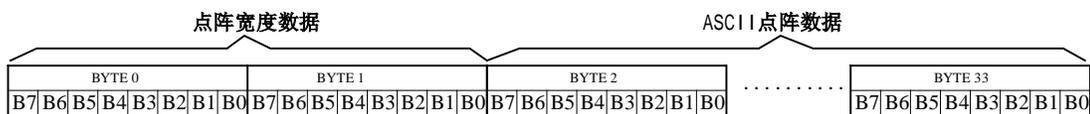


6.1.2 16 点阵不等宽阿拉伯字符排列格式

16 点阵不等宽阿拉伯字符的信息需要 34 个字节（BYTE 0 – BYTE33）来表示。

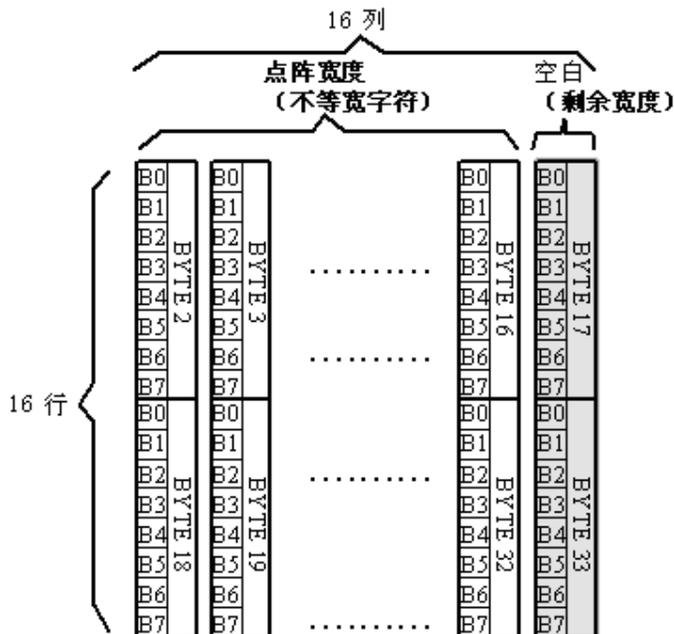
n 存储格式

由于字符是不等宽的，因此在存储格式中 BYTE0~ BYTE1 存放点阵宽度数据，BYTE2-33 存放竖置横排点阵数据。具体格式见下图：



n 存储结构

不等宽字符的点阵存储宽度是以 BYTE 为单位取整的，根据不同字符宽度会出现相应的空白区。根据 BYTE0~ BYTE1 所存放点阵的实际宽度数据，可以对还原下一个字的显示或排版留作参考。



例如：ASCII 方头字符 B

0-33BYTE 的点阵数据是： 00 0C 00 F8 F8 18 18 18 18 18 F8 F0 00 00 00 00 00 00 00 7F 7F 63 63 63 63 67 3E 1C 00 00 00 00 00

其中：

BYTE0~ BYTE1: 00 0C 为 ASCII 方头字符 B 的点阵宽度数据，即：12 位宽度。字符后面有 4 位空白区，可以在排版下一个字时考虑到这一点，将下一个字的起始位置前移。

BYTE2-33: 00 F8 F8 18 18 18 18 18 F8 F0 00 00 00 00 00 00 00 7F 7F 63 63 63 63 67 3E 1C 00 00 00 00 00 为 ASCII 方头字符 B 的点阵数据。

6.2 点阵字库地址表

序号	分类	字库内容	编码体系 (字符集)	字符数	容量	起始地址
1	Unicode 多国字符	8X16 点拉丁文系字符	Unicode	496	7936	0
2		8X16 点希腊文系字符	Unicode	96	1536	
3		8X16 点西里尔文系字符	Unicode	208	3328	
4		8X16 点希伯来文系字符	Unicode	112	1792	
5		8X16 点泰文系字符	Unicode	128	2048	
6		16 点阵不等宽阿拉伯文系字符	Unicode	576	19584	16640
7	汉字区	UNICODE 转换表	Unicode	20902	41804	45632
8		GB2312 转换表	GB2312	3008	6016	87436
9		16 点阵字符区	工作区	294	9408	36224
10		16 点阵汉字区	工作区	5270	168672	93452

6.3 字符在芯片中的地址计算方法

用户只要知道字符的内码，就可以计算出该字符点阵在芯片中的地址，然后就可从该地址连续读出点阵信息用于显示。

6.3.1 Unicode 多国字符的地址计算

6.3.1.1 8X16 等宽 Unicode 字符

包括 拉丁文、希腊文、基里尔文、希伯来文、泰文
说明：

BaseAdd: 说明本套字库在字库芯片中的起始字节地址。

unicode: 表示 unicode 内码 (16bits)

Address: 表示汉字点阵在芯片中的字节地址。

计算方法:

BaseAdd = 0;

if(unicode >= 0x20 && unicode <= 0x7f)//latin 1

Address = 16*(unicode-0x20)+ BaseAdd;

else if(unicode >= 0xa0 && unicode <= 0xff)//latin 2

Address = 16*(96+unicode-0xa0)+ BaseAdd;

else if(unicode >= 0x100 && unicode <= 0x17f)//latin 3

Address = 16*(96+96+unicode-0x100)+ BaseAdd;

else if(unicode >= 0x1a0 && unicode <= 0x1cf)//latin 4

Address = 16*(128+96+96+unicode-0x1a0)+ BaseAdd;

else if(unicode >= 0x1f0 && unicode <= 0x1ff)//latin 5

Address = 16*(48+128+96+96+unicode-0x1f0)+ BaseAdd;

else if(unicode >= 0x210 && unicode <= 0x21f)//latin 6

Address = 16*(16+48+128+96+96+unicode-0x210)+ BaseAdd;

else if(unicode >= 0x1ea0 && unicode <= 0x1eff)//latin 7

Address = 16*(16+16+48+128+96+96+unicode-0x1ea0)+ BaseAdd;

else if(unicode >= 0x370 && unicode <= 0x3cf)//greek

Address = 16*(96+16+16+48+128+96+96+unicode-0x370)+ BaseAdd;

else if(unicode >= 0x400 && unicode <= 0x45f)//cyrillic 1

```

        Address = 16*(96+96+16+16+48+128+96+96+unicode-0x400)+ BaseAdd;
    else if( unicode >= 0x490 && unicode <= 0x4ff )//cyrillic 2
        Address = 16*(96+96+96+16+16+48+128+96+96+unicode-0x490)+ BaseAdd;
    else if( unicode >= 0x590 && unicode <= 0x5ff )//hebrew
        Address = 16*(112+96+96+96+16+16+48+128+96+96+unicode-0x590)+ BaseAdd;
    else if( unicode >= 0xe00 && unicode <= 0xe7f )//thai
        Address = 16*(112+112+96+96+96+16+16+48+128+96+96+unicode-0xe00)+ BaseAdd;
    else
        Address = BaseAdd;
    
```

6.3.1.2 16 点阵不等宽阿拉伯文系字符

说明:

BaseAdd: 说明本套字库在字库芯片中的起始字节地址。

unicode_alb: 表示 unicode 内码 (16bits)

Address: 表示汉字点阵在芯片中的字节地址。

计算方法:

BaseAdd= 16640;

if(unicode_alb >= 0x0600 && unicode_alb <= 0x06FF)//alb 1

 Address = 34*(unicode_alb-0x0600)+ BaseAdd;

else if(unicode_alb >= 0xfb50 && unicode_alb <= 0xfbff)//alb 2

 Address = 34*(16*16+unicode_alb-0xfb50)+ BaseAdd;

else if(unicode_alb >= 0xfe70 && unicode_alb <= 0xfeff)//alb 3

 Address = 34*(16*11+16*16+unicode_alb-0xfe70)+ BaseAdd;

6.3.2 汉字的地址计算

6.3.2.1 15X16 GB2312 编码汉字字符

ZFAdd、HZAdd: 表示字符和汉字的起始地址。

MSB、LSB: 表示 GB2312 编码的高字节和低字节。

GB2312tab: 表示 GB2312 转码表起始地址。

Address: 表示汉字点阵在芯片中的字节地址。

ZFAdd = 36224;

HZAdd = 93452;

GB2312tab = 87436;

Unsigned char buffer[2];

if(MSB >= 0xA1 && MSB <= 0xA5)//字符区

{

if(MSB == 0xA1 && LSB >= 0xA1 && LSB <= 0xBF)

Address = 32*(LSB-0xA1)+ ZFAdd;

else if(MSB == 0xA3 && LSB >= 0xA1 && LSB <= 0xFE)

Address = 32*(31+LSB-0xA1)+ ZFAdd;

else if(MSB == 0xA4 && LSB >= 0xA1 && LSB <= 0xF3)

Address = 32*(31+94+LSB-0xA1)+ ZFAdd;

else if(MSB == 0xA5 && LSB >= 0xA1 && LSB <= 0xF6)

Address = 32*(31+94+83+LSB-0xA1)+ ZFAdd;

else

Address = ZFAdd;

}

else if((MSB >= 0xB0 && MSB <= 0xD7)&&(LSB>=0xA1 && LSB <= 0xFE))//汉字 5270

Address = 32*((MSB - 0xB0) * 94 + (LSB - 0xA1)+1)+ HZAdd;

else if((MSB >= 0xD8 && MSB <= 0xF7)&&(LSB>=0xA1 && LSB <= 0xFE))//汉字 5270~6763

{

Address = ((MSB - 0xD8) * 94 + (LSB - 0xA1)) *2 + GB2312tab;

GT_Read(buffer, Address,2); //从字库中读取二个字节

Address = 32* (256*buffer[0]+buffer[1]) + HZAdd;

}

6.3.2.2 15X16 Unicode 编码汉字字符

ZFAdd、HZAdd: 表示字符和汉字的起始地址。

Unicode: 表示汉字的 unicode 编码。

Unicodetab: 表示 unicode 转码表起始地址。

Address: 表示汉字点阵在芯片中的字节地址。

ZFAdd = 36224;

HZAdd = 93452;

Unicodetab = 45632;

Unsigned char buffer[2];

```

if(Unicode >= 0x4e00 && Unicode < 0x9fa6 )    // Unicode 汉字区
{
    Address = 2*(Unicode-0x4e00);
    Address = Address + Unicodetab;
    GT_Read(buffer, Address,2); //从字库中读取二个字节
    Address = 256*buffer[0]+buffer[1];
    Address = 32* Address + HZAdd;
}
else if(Unicode >= 0xff01 && Unicode <= 0xff5d )    //Unicode 字符区 1
    Address = 32*(31+Unicode-0xff01)+ ZFAdd;
else if(Unicode == 0xffe3)                        //Unicode 字符区 2
    Address = 32*(31+93)+ ZFAdd;
else if( Unicode >= 0x3041 && Unicode <= 0x3096 ) //Unicode 字符区 3
    Address = 32*(31+93+1+Unicode-0x3041)+ ZFAdd;
else if( Unicode >= 0x30a1 && Unicode <= 0x30f6 ) //Unicode 字符区 4
    Address = 32*(31+93+1+83+Unicode-0x30a1)+ ZFAdd;
else                                              //Unicode 字符区 5
    Address = 32* find_zf_tab(Unicode)+ ZFAdd;
    
```

其中函数 unsigned char find_zf_tab(unsigned short unicode) 参考下列代码

```

unsigned short unicode_zf[31] =
{
    0x3000,0x3001,0x3002,0x00b7,0x02c9,0x02c7,0x00a8,0x3003,0x3005,0x2014,0xff5e,0x2016,0x2026,
    0x2018,0x2019,0x201c,0x201d,0x3014,0x3015,0x3008,0x3009,0x300a,0x300b,0x300c,0x300d,0x300e
    ,0x300f,0x3016,0x3017,0x3010,0x3011
};
unsigned char find_zf_tab(unsigned short unicode)
{
    unsigned char i;
    for( i=0; i < 31; i++ )
    {
        if( unicode == unicode_zf[i] )
            break;
    }
    if(i==31)
        return(0);
    return(i);
}
    
```

7 附录

7.1 Unicode 字符区字符

Unicode 字符区共收录 拉丁文系 (LATIN)、希腊文系 (GREEK)、基里尔文系 (CYRILLIC)、希伯来文系 (HEBREW)、泰文 (THAI) 阿拉伯文系 (ARABIC)。

7.1.1 8×16 点拉丁文系 (496 字符)

Unicode 字符区-拉丁文系共计收录 496 个字符。

说明：由于拉丁文系 (BASIC LATIN) 的内码范围 0020~007E 与基本 ASCII 编码共用。

Unicode 字符区-拉丁文系

00	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
2		!	"	#	\$	%	&	'	()	*	+	,	-	.	/
3	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
6	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7	P	q	r	s	t	u	v	w	x	y	z	{		}	~	

00	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
A		ı	ç	£	¤	¥	ı	§	¨	©	ª	«	¬	-	®	¯
B	°	±	²	³	´	µ	¶	·	¸	¹	º	»	¼	½	¾	¿
C	À	Á	Â	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ì	Í	Î	Ï
D	Ð	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü	Ý	Þ	ß
E	à	á	â	ã	ä	å	æ	ç	è	é	ê	ë	ì	í	î	ï
F	ð	ñ	ò	ó	ô	õ	ö	÷	ø	ù	ú	û	ü	ý	þ	ÿ

Unicode 字符区-拉丁文系

01	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	Ā ā	Ă ă	Ą ą	Ć ć	Ĉ ĉ	Ċ ċ	Č č	Ď ě	Đ đ							
1	Ē ē	Ĕ ĕ	Ė ė	Ę ę	Ě ě	Ĝ ğ	Ğ ğ	Ģ ģ								
2	Ĥ ĥ	Ħ ħ	Ĩ ĩ	Ī ī	Ĵ ĵ	Ķ ķ	Ŀ ź	Ł ł	Ń ń	Ņ ņ						
3	Ŧ ŧ	Ũ ů	Ū ū	Ŵ ŵ	Ŷ ŷ	Ÿ Ź	Ż ż	Ž ž								
4	ı	Ł	ł	Ń	ń	Ņ	ņ	Ň	ň	Ŋ	ŋ	Ō	ō	Ŏ	ö	
5	Œ	œ	Ŕ	ŕ	Ŗ	ŗ	Ř	ř	Ś	ś	Ŝ	ŝ	Ş	ş	Ş	ş
6	Š	š	Ţ	ţ	Ț	ț	Ț	ț	Ț	ț	Ț	ț	Ț	ț	Ț	ț
7	Ț	ț	Ț	ț	Ț	ț	Ț	ț	Ț	ț	Ț	ț	Ț	ț	Ț	ț

01	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
A	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ
B	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ
C	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ
F	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ

02	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
1	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ

1E	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
A	Ā	ā	Ă	ă	Ą	ą	Ć	ć	Ĉ	ĉ	Ċ	ċ	Č	č	Ď	ď
B	Ē	ē	Ĕ	ĕ	Ė	ė	Ę	ę	Ě	ě	Ĝ	ğ	Ğ	ğ	Ģ	ģ
C	Ĥ	ĥ	Ħ	ħ	Ĩ	ĩ	Ī	ī	Ĵ	ĵ	Ķ	ķ	Ŀ	ź	Ł	ł
D	Ŧ	ŧ	Ũ	ů	Ū	ū	Ŵ	ŵ	Ŷ	ŷ	Ÿ	Ź	Ż	ż	Ž	ž
E	ı	Ł	ł	Ń	ń	Ņ	ņ	Ň	ň	Ŋ	ŋ	Ō	ō	Ŏ	ö	
F	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ	Œ	œ

7.1.2 8×16 点希腊文系（96 字符）

Unicode 字符区-希腊文系共计收录 96 个字符。

Unicode 字符区-希腊文系

03	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
7					'	,					,	α	ε	ε	,	
8					'	..	Α	·	Ε	Η	Ι		Ο		Τ	Ω
9	ι	Α	Β	Γ	Δ	Ε	Ζ	Η	Θ	Ι	Κ	Λ	Μ	Ν	Ξ	Ο
A	Π	Ρ		Σ	Τ	Υ	Φ	Χ	Ψ	Ω	Ϊ	Ϋ	Ό	Έ	Ή	Ί
B	ϐ	α	β	γ	δ	ε	ζ	η	θ	ι	κ	λ	μ	ν	ξ	ο
C	π	ρ	ς	σ	τ	υ	φ	χ	ψ	ω	ϊ	ϋ	ό	ύ	ώ	

7.1.3 8X16 点基里尔文系（208 字符）

Unicode 字符区-基里尔文系共计收录 208 个字符。

Unicode 字符区-基里尔文系

04	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	Ё	ѐ	Ѣ	ѣ	Є	є	І	і	Ј	ѣ	Ѥ	ѥ	Ѧ	ѧ	Ѩ	ѩ
1	А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П
2	Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я
3	а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п
4	р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я
5	ё	ѓ	ѣ	ѣ	є	є	і	і	ј	ѣ	ѥ	ѥ	Ѧ	ѧ	Ѩ	ѩ

04	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
9	Г	г	Ғ	ғ	Б	б	Ж	ж	З	з	К	к	Қ	қ	К	к
A	К	к	Ң	ң	Н	н	Ь	ь	Ѡ	ѡ	С	с	Т	т	У	у
B	Ү	ү	Х	х	Ц	ц	Ч	ч	Ч	ч	Ң	ң	Ѣ	ѣ	Ѣ	ѣ
C	І	ӡ	Ӣ	ӣ	Ӥ	ӥ	Ӧ	ӧ	Ө	ө	Ч	ч	Ч	ч	М	м
D	Ӓ	ӓ	Ӕ	ӕ	Ӗ	ӗ	Ә	ә	Ӛ	ӛ	Ӝ	ӝ	Ӟ	ӟ	Ӡ	ӡ
E	Ӣ	ӣ	Ӥ	ӥ	Ӧ	ӧ	Ә	ә	Ӛ	ӛ	Ӝ	ӝ	Ӟ	ӟ	Ӡ	ӡ
F	Ӣ	ӣ	Ӥ	ӥ	Ӧ	ӧ	Ә	ә	Ӛ	ӛ	Ӝ	ӝ	Ӟ	ӟ	Ӡ	ӡ

7.1.4 8X16 点希伯来文系（112 字符）

Unicode 字符区-希伯来文系共计收录 112 个字符。

05	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
9		א	ב	ג	ד	ה	ו	ז	ח	ט	י	כ	ל	מ	נ	ס
A	פ	צ	ק	ר	ש	ת										
B	׀	ׁ	ׂ	׃	ׄ	ׅ	׆	ׇ	׈	׉	׊	׋	׌	׍	׎	׏
C	א	ב	ג	ד	ה	ו	ז	ח	ט	י	ך	כ	ל	ם	מ	ן
D	נ	ס	ע	ף	פ	ץ	צ	ק	ר	ש	ת	׫	׬	׭	׮	ׯ
E	װ	ױ	ײ	׳	״	׵	׶	׷	׸	׹	׺	׻	׼	׽	׾	׿
F	׿	׿	׿	׿	׿	׿	׿	׿	׿	׿	׿	׿	׿	׿	׿	׿

7.1.5 8X16 点泰文文系（128 字符）

Unicode 字符区-泰文文系共计收录 128 个字符。

Thai

0E	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0		ก	ข	ฃ	ค	ฅ	ง	จ	ฉ	ช	ฌ	ญ	ฎ	ฏ	ฐ	ฑ
1	ฒ	ณ	ด	ต	ถ	ท	ธ	น	บ	ป	ผ	ฝ	พ	ฟ	ภ	ผ
2	ภ	ม	ย	ร	ล	ว	ศ	ษ	ส	ห	ฬ	อ	ฮ	า	ิ	ี
3	ุ	ู	เ	แ	อ	ก	ข	ฃ	ค	ฅ	ง	จ	ฉ	ช	ฌ	ญ
4	ฎ	ฏ	ฐ	ฑ	ฒ	ณ	ด	ต	ถ	ท	ธ	น	บ	ป	ผ	ฝ
5	พ	ฟ	ภ	ผ	ฐ	ฑ	ฒ	ณ	ด	ต	ถ	ท	ธ	น	บ	ป
6																
7																

7.1.6 16 点阵阿拉伯文系

Unicode 阿拉伯文系编码字符

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
060													،			
061												؛				؟
062		ء	آ	أ	ؤ	إ	ئ	ا	ب	ة	ت	ث	ج	ح	خ	د
063	ذ	ر	ز	س	ش	ص	ض	ط	ظ	ع	غ					
064	-	ف	ق	ك	ل	م	ن	هـ	و	ى	ي	ٴ	ٶ	ٷ	ٸ	ٹ
065	ٺ	ٻ	ټ	ٽ	ټ	ٺ										
066	ٲ	ٱ	ٲ	ٳ	ٴ	ٵ	ٶ	ٷ	ٸ	ٹ	٪	،	،	*		
067	'	آ	أ	إ	ٴ	ٶ	ؤ	ى	ٺ	ت	ب	ټ	ت	ټ	ت	ټ

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
068	پ	خ	خ	ج	ج	خ	ج	چ	ڈ	د	د	ڈ	ذ	د	ذ	ذ
069	ڈ	ڑ	ڑ	ر	ر	ر	ر	ز	ژ	ژ	س	س	س	ص	ض	ض
06A	غ	ف	ف	ف	ف	ف	ق	ق	ق	ق	ک	ک	ک	ک	ک	ک
06B	گ	گ	گ	گ	گ	ل	ل	ل	ل	ن	ن	ن	ن	ن	ن	ن
06C	ة	هـ	هـ	هـ	و	و	و	و	و	و	و	و	و	و	و	و
06D	ي	ي	ے	ے	-	ٴ	ٶ	ٷ	ٸ	ٹ	ٺ	ٻ	ټ	ٽ	ټ	ٺ
06E	ٲ	ٱ	ٲ	ٳ	ٴ	ٵ	ٶ	ٷ	ٸ	ٹ	٪	،	،	*		
06F	ٲ	ٱ	ٲ	ٳ	ٴ	ٵ	ٶ	ٷ	ٸ	ٹ						

Unicode 阿拉伯文系编码字符

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
FB5	آ	أ	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب
FB6	ز	ز	ث	ث	ث	ث	ث	ث	ث	ث	ث	ث	ث	ث	ث	ث
FB7	ق	ق	ح	ح	ح	ح	ح	ح	ح	ح	ح	ح	ح	ح	ح	ح

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
FB8	چ	چ	د	د	د	د	د	د	د	د	د	د	د	د	د	د
FB9	ک	ک	گ	گ	گ	گ	گ	گ	گ	گ	گ	گ	گ	گ	گ	گ
FBA	ٹ	ٹ	ط	ط	ء	ء	ء	ء	ء	ء	ء	ء	ء	ء	ء	ء
FBB	ئ	ئ														
FBC																
FBD			ك	ك	ك	ك	ق	ق	ق	ق	ق	ق	ق	ق	ق	ق
FBE	و	و	و	و	ي	ي	ي	ي	ي	ي	ي	ي	ي	ي	ي	ي
FBF	ع	ع	ع	ع	ع	ع	ي	ي	ي	ي	ي	ي	ي	ي	ي	ي

	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
FE7	°	°	°		°		°	°	°	°	°	°	°	°	°	°
FE8	ء	أ	أ	أ	أ	أ	أ	أ	أ	أ	أ	أ	أ	أ	أ	أ
FE9	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب	ب
FEA	ج	ح	ح	ح	ح	ح	ح	ح	ح	ح	ح	ح	ح	ح	ح	ح
FEB	ز	س	س	س	س	س	س	س	س	س	س	س	س	س	س	س
FEC	ط	ط	ط	ط	ط	ط	ط	ط	ط	ط	ط	ط	ط	ط	ط	ط
FED	ق	ق	ق	ق	ق	ق	ق	ق	ق	ق	ق	ق	ق	ق	ق	ق
FEE	ك	م	م	م	م	ن	ن	ن	ن	ن	ن	ن	ن	ن	ن	ن
FEF	ي	ي	ي	ي	ي	ل	ل	ل	ل	ل	ل	ل	ل			

7.2 175 国中外文字库汉语拼音索引表

175 国中外文字库汉语拼音索引表

首字母	国家	总表序号	首字母	国家	总表序号	首字母	国家	总表序号
A	阿富汗	171		哥斯达黎加	67	P	葡萄牙	86
	阿尔巴尼亚	122		圭亚那	15		帕劳	20
	阿尔及利亚	152	H	海地	44	R	瑞典	102
	阿曼	164		洪都拉斯	71		瑞士	94
	阿根廷	78		荷兰	98	S	萨尔瓦多	70
	阿联酋	158		哈萨克斯坦	144		圣马力诺	118
	阿塞拜疆	139	J	吉尔吉斯斯坦	140		圣多美和普林西比	90
	爱尔兰	2		吉布提	155		圣基茨—尼维斯	14
	爱沙尼亚	107		吉布提	56		圣文森特	11
	奥地利	95		津巴布韦	27		圣卢西亚岛	12
	澳大利亚	16		加蓬	60		塞浦路斯	147
	安道尔	64		加拿大	4		塞内加尔	45
	安提瓜和巴布达	9		加纳	31		塞拉利昂	29
	安哥拉	91		几内亚	48		塞舌尔	36
	埃及	148		几内亚比绍	89		索马里	154
B	巴拿马	68		基里巴斯	23		苏丹	153
	巴黎嫩	159		捷克	110		沙特阿拉伯	167
	巴林	163	K	喀麦隆	53		所罗门	21
	巴基斯坦	170		科摩罗	61		斯洛伐克	111
	巴勒斯坦	168		科特迪瓦	49		斯洛文尼亚	115
	巴拉圭	82		卡塔尔	162		苏里南	99
	巴哈马	8		肯尼亚	129	T	土耳其	120
	巴西	87		克罗地亚	116		土库曼斯坦	142
	白俄罗斯	132		科威特	161		突尼斯	149
	冰岛	105	L	利比里亚	30		特立尼达和多巴哥	7
	贝宁	51		利比亚	150		汤加	18
	伯利兹	5		拉脱维亚	108		塔吉克斯坦	141
	比利时	42		立陶宛	109		泰国	173
	波罗黎各	73		罗马尼亚	114		坦桑尼亚	130
	玻利维亚	83		列支敦士登	97	W	瓦努阿图	22
	波兰	112		卢森堡	96		危地马拉	66
	布隆迪	57		莱索托	40		文莱	125
	波黑	137	M	马耳他	121		乌拉圭	81
	保加利亚	134		马其顿	138		乌克兰	133
	布基纳法索	47		马达加斯加	62		乌干达	33
	博茨瓦纳	38		马来西亚	124		乌兹别克斯坦	143
C	赤道新几内亚	84		马拉维	35		委内瑞拉	75
D	德国	93		马里	46	X	西班牙	63
	丹麦	100		马绍尔群岛	25		休达和梅利亚	85
	多米尼加	10		摩尔多瓦	135		希腊	146
	多米尼加共和国	69		摩纳哥	43		叙利亚	157

	东帝汶	127	摩洛哥	151	新西兰	17
	多哥	50	毛里求斯	37	新加坡	175
E	俄罗斯	131	毛里塔尼亚	156	匈牙利	113
	厄瓜多尔	79	秘鲁	77	Y 牙买加	6
F	佛得角	88	美国	3	也门	160
	梵提冈	119	蒙古	145	意大利	117
	法罗群岛	103	墨西哥	65	约旦	165
	芬兰	106	民主刚果	58	印度尼西亚	126
	菲律宾	128	莫桑比克	92	越南	123
	法国	41	N 纳米比亚	39	伊拉克	166
	斐济	19	南非	26	伊朗	169
G	冈比亚	28	瑙鲁	24	英国	1
	古巴	74	尼加拉瓜	72	以色列	172
	哥伦比亚	76	尼日利亚	32	Z 赞比亚	34
	格林纳达	13	尼日尔	52	中非	55
	格陵兰	104	南斯拉夫联盟	136	中国	174
	刚果	59	挪威	101	乍得	54
					智利	80

7.3 175 国中外文字库英文字母索引表

175 国中外文字库英文字母索引表

首字母	国家	总表序号	首字母	国家	总表序号	首字母	国家	总表序号
A	Afghanistan	171	G	Gambia	28	P	Pakistan	170
	Albania	122		Germany	93		Palau	20
	Algeria	152		Ghana	31		Palestine	168
	Andorra	64		Greece	146		Panama	68
	Angola	91		Greenland	104		Paraguay	82
	Antigua and Barbuda	9		Grenada	13		Peru	77
	Argentina	78		Guatemala	66		Philippines, The	128
	Australia	16		Guinea	48		Poland	112
	Austria	95		Guinea-Bissau	89		Portugal	86
	Azerbaijan	139		Guyana	15		Puerto Rico	73
B	Bahamas	8	H	Haiti	44	Q	Qatar	162
	Bahrain	163		Holland	98	R	Republic of Democratic Congo	58
	Barbados	137		Honduras	71		Romania	114
	Belgium	42		Hungary	113		Russia	131
	Belize	5	I	Iceland	105	S	San Marino	118
	Benin	51		Indonesia	126		Sao Tome and Principe	90
	Bolivia	83		Iran	169		Saudi Arabia	167
	Botswana	38		Iraq	166		Senegal	45
	Brazil	87		Israel	172		Seychelles	36
	Britain 或 United Kingdom	1		Italy	117		Sierra Leone	29
	Brunei	125	J	Jamaica	6		Singapore	175
	Bulgaria	134		Jordan	165		Slovakia	111
	Burkina Faso	47	K	Kazakhstan	144		Slovenia	115
	Burundi	57		Kenya	129		Solomon	21
	Byelorussia 或 Belarus	132		Kirghizstan	140		Somalia	154
C	Cameroon	53		Kiribati	23		South Africa	26
	Canada	4		Kuwait	161		Spain	63
	Cape Verde	88	L	Latvia	108		St.Kitts-Nevis	14
	Central African Republic	55		Lebanon	159		St.Lucia	12
	Ceuta and Melilla	85		Lesotho	40		St.Vincent	11
	Chad	54		Liberia	30		Sudan, The	153
	Chile	80		Libya	150		Surinam	99
	China	174		Liechtenstein	97		Sweden	102
	Colombia	76		Lithuania	109		Switzerland	94
	Comoros	61		Lreland	2		Syria	157
	Congo	59		Luxembourg	96	T	Tajikistan	141
	Costa Rica	67	M	Macedonia	138		Tanzania	130
	cote dlvoire	49		Madagascar	62		Thailand	173
	Croatia	116		Malawi	35		Togo	50

	Cuba	74	Malaysia	124	Tonga	18
	Cyprus	147	Mali	46	Trinidad and Tobago	7
	Czech	110	Malta	121	Tunisia	149
D	Denmark	100	Marshall Islands	25	Turkey	120
	Djibouti	56	Mauritania	156	Turkmenistan	142
	Djibouti	155	Mauritius	37	U Uganda	33
	Dominica	10	Mexico	65	Ukraine	133
	Dominican Republic	69	Moldova	135	United Arab Emirates, The	158
E	East Timor	127	Monaco	43	Uruguay	81
	Ecuador	79	Mongolia	145	USA	3
	Egypt	148	Morocco	151	Uzbekistan	143
	El Salvador	70	Mozambique	92	V Vanuatu	22
	Equatorial New Guinea	84	N Namibia	39	Vatican	119
	Estonia	107	Nauru	24	Venezuela	75
F	F.R. Yugoslavia	136	New Zealand	17	Vietnam	123
	Faroes, The	103	Nicaragua	72	Y Yemen	160
	Fiji	19	Niger	52	Z Zambia	34
	Finland	106	Nigeria	32	Zimbabwe	27
	France	41	Norway	101		
G	Gabon	60	O Oman	164		

7.4 175 国中外文字库总表

文系	区域	序号	国家		语言
拉丁文 (英语)	欧洲	1	英国	Britain 或 United Kingdom	英语
		2	爱尔兰	Ireland	
	北美洲	3	美国	USA	英语
		4	加拿大	Canada	英语、法语
		5	伯利兹	Belize	英语
		6	牙买加	Jamaica	
		7	特立尼达和多巴哥	Trinidad and Tobago	
		8	巴哈马	Bahamas	
		9	安提瓜和巴布达	Antigua and Barbuda	
		10	多米尼加	Dominica	
		11	圣文森特	St. Vincent	
		12	圣卢西亚岛	St. Lucia	
		13	格林纳达	Grenada	
		14	圣基茨-尼维斯	St. Kitts-Nevis	
	南美洲	15	圭亚那	Guyana	英语
	大洋洲	16	澳大利亚	Australia	英语
		17	新西兰	New Zealand	
		18	汤加	Tonga	
		19	斐济	Fiji	
		20	帕劳	Palau	
		21	所罗门	Solomon	
		22	瓦努阿图	Vanuatu	
		23	基里巴斯	Kiribati	
		24	瑙鲁	Nauru	
		25	马绍尔群岛	Marshall Islands	
	非洲	26	南非	South Africa	英语、南非荷兰语
		27	津巴布韦	Zimbabwe	英语
		28	冈比亚	Gambia	
		29	塞拉利昂	Sierra Leone	
		30	利比里亚	Liberia	
		31	加纳	Ghana	
		32	尼日利亚	Nigeria	
		33	乌干达	Uganda	
		34	赞比亚	Zambia	
		35	马拉维	Malawi	
		36	塞舌尔	Seychelles	
		37	毛里求斯	Mauritius	
		38	博茨瓦纳	Botswana	
		39	纳米比亚	Namibia	
		40	莱索托	Lesotho	
拉丁文 (法语)	欧洲	41	法国	France	法语
		42	比利时	Belgium	法语、荷兰语
		43	摩纳哥	Monaco	法语、意大利语

	北美洲	44	海地	Haiti	法语
	非洲	45	塞内加尔	Senegal	法语 法语
		46	马里	Mali	
		47	布基纳法索	Burkina Faso	
		48	几内亚	Guinea	
		49	科特迪瓦	cote d'Ivoire	
		50	多哥	Togo	法语
		51	贝宁	Benin	
		52	尼日尔	Niger	
		53	喀麦隆	Cameroon	
		54	乍得	Chad	
55	中非	Central African Republic			
拉丁文 (法语)	非洲	56	吉布提	Djibouti	
		57	布隆迪	Burundi	
		58	民主刚果	Republic of Democratic Congo	
		59	刚果	Congo	法语
		60	加蓬	Gabon	
		61	科摩罗	Comoros	
62	马达加斯加	Madagascar			
拉丁文 (西班牙语)	欧洲	63	西班牙	Spain	西班牙语、加泰隆语
		64	安道尔	Andorra	西班牙语
	北美洲	65	墨西哥	Mexico	西班牙语
		66	危地马拉	Guatemala	
		67	哥斯达黎加	Costa Rica	
		68	巴拿马	Panama	
		69	多米尼加共和国	Dominican Republic	
		70	萨尔瓦多	El Salvador	
		71	洪都拉斯	Honduras	
		72	尼加拉瓜	Nicaragua	
		73	波多黎各	Puerto Rico	
	74	古巴	Cuba		
	南美洲	75	委内瑞拉	Venezuela	西班牙语
		76	哥伦比亚	Colombia	
		77	秘鲁	Peru	
		78	阿根廷	Argentina	
		79	厄瓜多尔	Ecuador	
		80	智利	Chile	
		81	乌拉圭	Uruguay	
82		巴拉圭	Paraguay		
83		玻利维亚	Bolivia		
非洲	84	赤道新几内亚	Equatorial New Guinea	西班牙语	
	85	休达和梅利利亚	Ceuta and Melilla		
拉丁文 (葡萄牙语)	欧洲	86	葡萄牙	Portugal	葡萄牙语
	南美洲	87	巴西	Brazil	
	非洲	88	佛得角	Cape Verde	
		89	几内亚比绍	Guinea-Bissau	

		90	圣多美和普林西比	Sao Tome and Principe	
		91	安哥拉	Angola	
		92	莫桑比克	Mozambique	
拉丁文 (德语)	欧洲	93	德国	Germany	德语
		94	瑞士	Switzerland	德语、法语
		95	奥地利	Austria	德语
		96	卢森堡	Luxembourg	德语、法语
		97	列支敦士登	Liechtenstein	德语
拉丁文 (荷兰语)	欧洲	98	荷兰	Holland	荷兰语
	南美洲	99	苏里南	Surinam	
拉丁文 (北欧)	欧洲	100	丹麦	Denmark	丹麦语
		101	挪威	Norway	挪威语
		102	瑞典	Sweden	瑞典语
		103	法罗群岛	Faroes, The	法罗语
		104	格陵兰	Greenland	格陵兰语
		105	冰岛	Iceland	冰岛语
		106	芬兰	Finland	芬兰语、瑞典语
		107	爱沙尼亚	Estonia	爱沙尼亚语
		108	拉脱维亚	Latvia	拉脱维亚语
		109	立陶宛	Lithuania	立陶宛语
拉丁文 (中欧)	欧洲	110	捷克	Czech	捷克语
		111	斯洛伐克	Slovakia	斯洛伐克语
		112	波兰	Poland	波兰语
		113	匈牙利	Hungary	匈牙利语
		114	罗马尼亚	Romania	罗马尼亚语
拉丁文 (中欧)	欧洲	115	斯洛文尼亚	Slovenia	斯洛文尼亚语
		116	克罗地亚	Croatia	克罗地亚语
拉丁文 (南欧)	欧洲	117	意大利	Italy	意大利语
		118	圣马力诺	San Marino	
		119	梵提冈	Vatican	
		120	土耳其	Turkey	土耳其语
		121	马耳他	Malta	马耳他语
		122	阿尔巴尼亚	Albania	阿尔巴尼亚语
拉丁文 (东南亚)	亚洲	123	越南	Vietnam	越南语
		124	马来西亚	Malaysia	马来语
		125	文莱	Brunei	
		126	印度尼西亚	Indonesia	印尼语
		127	东帝汶	East Timor	
		128	菲律宾	Philippines, The	英语、塔加洛语
拉丁文 (非洲)	非洲	129	肯尼亚	Kenya	斯瓦西里语
		130	坦桑尼亚	Tanzania	
西里尔文 (东欧)	欧洲	131	俄罗斯	Russia	俄语
		132	白俄罗斯	Byelorussia 或 Belarus	
		133	乌克兰	Ukraine	俄语、乌克兰语
		134	保加利亚	Bulgaria	保加利亚语
		135	摩尔多瓦	Moldova	俄语
		136	南斯拉夫联盟	F.R.Yugoslavia	塞尔维亚语

		137	波黑	Barbados	塞尔维亚语
		138	马其顿	Macedonia	马其顿语
西里尔文 (亚洲)	亚洲	139	阿塞拜疆	Azerbaijan	阿塞拜疆语
		140	吉尔吉斯斯坦	Kirghizstan	吉尔吉斯斯坦语
		141	塔吉克斯坦	Tajikistan	塔吉克斯坦语
		142	土库曼斯坦	Turkmenistan	土库曼斯坦语
		143	乌兹别克斯坦	Uzbekistan	乌兹别克斯坦语
		144	哈萨克斯坦	Kazakhstan	哈萨克斯坦语
		145	蒙古	Mongolia	蒙古语
希腊文	亚洲	146	希腊	Greece	希腊语
		147	塞浦路斯	Cyprus	
阿拉伯文 (非洲)	非洲	148	埃及	Egypt	阿拉伯语
		149	突尼斯	Tunisia	
		150	利比亚	Libya	
		151	摩洛哥	Morocco	
		152	阿尔及利亚	Algeria	
		153	苏丹	Sudan, The	
		154	索马里	Somalia	
		155	吉布提	Djibouti	
		156	毛里塔尼亚	Mauritania	
阿拉伯文 (亚洲)	亚洲	157	叙利亚	Syria	阿拉伯语
		158	阿联酋	United Arab Emirates, The	
		159	黎巴嫩	Lebanon	
		160	也门	Yemen	
		161	科威特	Kuwait	
		162	卡塔尔	Qatar	
		163	巴林	Bahrain	
		164	阿曼	Oman	
		165	约旦	Jordan	
		166	伊拉克	Iraq	
		167	沙特阿拉伯	Saudi Arabia	
		168	巴勒斯坦	Palestine	
		169	伊朗	Iran	
		170	巴基斯坦	Pakistan	乌尔都语、阿拉伯语
171	阿富汗	Afghanistan	普什图语		
希伯来文	亚洲	172	以色列	Israel	希伯来语
泰文	亚洲	173	泰国	Thailand	泰语
中文	亚洲	174	中国	China	汉语
		175	新加坡	Singapore	

7.5 UTF8 编码到 UTF16 编码转换程序

```
//-----
//函数名称: UFT8toUTF16(unsigned char *pUTF8,unsigned short *pUTF16 )
//函数说明: UTF8 码转 UTF16 码函数
//参数说明: pUTF8 UTF8 数据指针; pUTF16 获得 UTF16 的指针
//返回说明: 该 UTF8 字符的占用字节数。
//版权 上海集通数码科技有限公司 时间: 2010-11-22
//-----
unsigned char UFT8toUTF16(unsigned char *pUTF8,unsigned short *pUTF16 )
{
    unsigned char bytes[3];
    unsigned short unicode16;

    bytes[0] = *pUTF8++;
    if( bytes[0] < 0x80 )
    {
        *pUTF16 = bytes[0];
        return(1);
    }
    bytes[1] = *pUTF8++;
    if( bytes[0] >= 0xC0 && bytes[0] < 0xE0 )
    {
        unicode16 = 0x1f&bytes[0];
        *pUTF16 = (unicode16<<6)+(0x3f&bytes[1]);
        return(2);
    }
    bytes[2] = *pUTF8++;
    if( bytes[0] >= 0xE0 && bytes[0] < 0xF0 )
    {
        unicode16 = 0x0f&bytes[0];
        unicode16 = (unicode16<<6)+(0x3f&bytes[1]);
        *pUTF16 = (unicode16<<6)+(0x3f&bytes[2]);
        return(3);
    }
    else
        return(0);
}
}
```