

硬件汉字库的原理与应用

www.mcutech.cn

摘要 本文论述了硬件汉字库的基本原理和具体制作步骤。并利用 89C51 作为主控 CPU，通过 TFT 控制板来点 320240 液晶屏。给出了在液晶屏上显示“国”字过程。

关键词 硬件汉字库 液晶显示 89C51

随着液晶显示技术的发展和运用，越来越多的开发人员希望在自己开发的仪器设备中使用液晶屏幕来显示汉字。我们日常生活的一些手持工具上如手机快译通等等，它们可以任意输入汉字不仅灵活性大而且检索速度快，这其中的主要原因就是在这些设备中固化了硬件汉字库，越来越多的技术开发人员开始关注起硬件汉字库，希望能在自己开发的设备中设计并使用硬件汉字库，达到加快产品开发速度、降低软件成本的目的。

一 硬件汉字库相关知识

1 汉字机内码、国标码和区位码

在PC 机的文本文件中，汉字是以机内码的形式存储的，每个汉字占用两个字节长度，为了和ASCII码区别，范围从十六进制的0A1H 开始，小于80H 的为ASCII 码将机内码，每个字节的最高位屏蔽掉再以十六进制的形式显示出来则为国标码，将机内码的每个字节各减去0A0H 再以十进制显示出来即为该汉字的区位码。例如“国”字的机内码、国标码和区位码如表 1 所示。

类别	数值	高位字节								低位字节							
机内码	B9FAH	1	0	1	1	1	0	0	1	1	1	1	1	0	1	0	
国标码	397AH	0	0	1	1	1	0	0	1	0	1	1	1	0	1	0	
区位码	195AH	0	0	0	1	1	0	0	1	0	1	0	1	0	1	0	

即：区位码=机内码-0A0A0H； 就“国”字而言，其区位码和机内码的关系为：

195AH（区位码）= 0B9FAH（机内码）- 0A0A0H

记住这个关系是我们理解下面有关程序的关键。

2 国标汉字字符集与区位码

根据对汉字使用频率程度的研究，可把汉字分成高频字（约100个），常用字（约3000个），次常用字（约4000字），罕见字（约8000个）和死字（约45000个）。即正常使用的汉字达15000 个。我国1981年公布了《通讯用汉字字符集（基本集）及其交换码标准》GB2312-80方案。把高频字、常用字和次常用字集成汉字基本字符集共6763个，在该字符集中按汉字使用的频度又将其分为一级汉字3755 个按拼音排序，二级汉字3008个按部首排序，再加上西文字母数字图形符号等700个。国家标准的汉字字符集GB2312-80 在汉字操作系统中是以汉字库的形式提供的。汉字库结构作了统一规定如图 1 所示

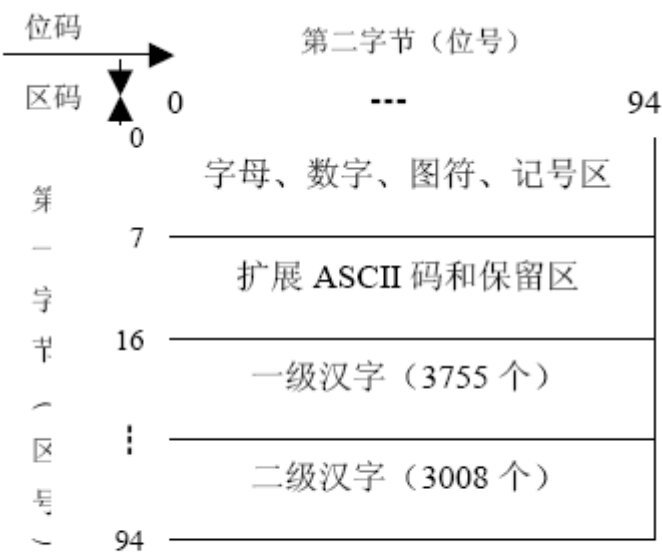


图1 国标（GB2312-80）汉字字符集

即将字库分成94 个区，每个区有94 个汉字以位作区别，每一个汉字在汉字库中有确定的区和位，编号用两个字节，这就是所谓的区位码。区位码的第一个字节表示区号，第二个字节表示位号，因而只要知道了区位码就可知道该汉字在字库中的地址，每个汉字在字库中是以点阵字模形式存储的，如一般采用16x 16 点阵形式，每个点用一个二进制位表示，存1 的点当显示时可以在屏上显示一个亮点，存0 的点则在屏上不显示，这样就把存某字的16x 16 点阵信息直接用来在显示器上按上述原则显示则将出现对应的汉字，如一个国字的16x 16 点阵字模如图2 所示，当用存储单元存储该字模信息时将需32 个字节地址，在图2 的右边写出了该字模对应的字节值。

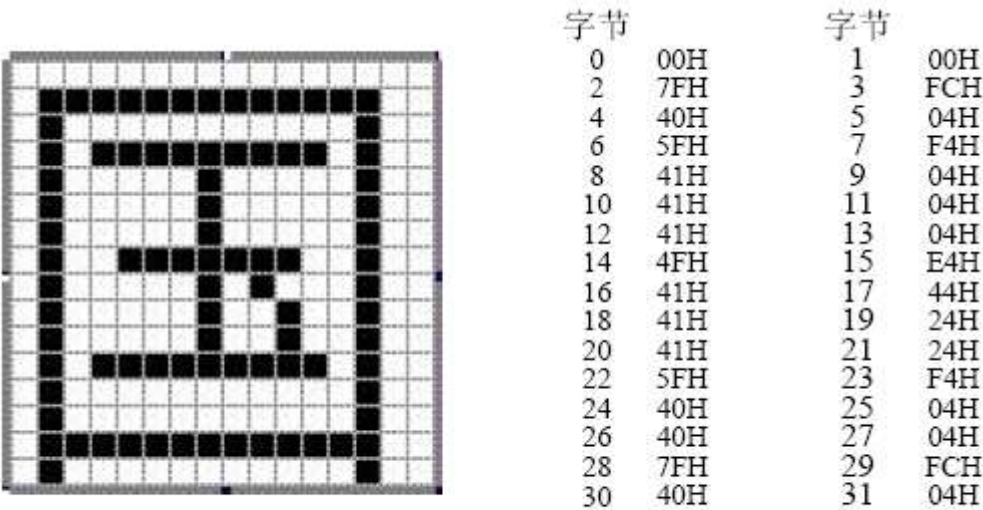


图2 “国”字的 16×16 点阵字模

二 汉字库的制作

在本节中笔者将论述将16x 16 点阵的宋体汉字库烧录到256KB 的存储器中的
武汉东湖新技术开发区珞瑜东路 76 号 电话：027-87680042 138-71489945 137-97051554

www.mcutech.cn

具体方法，这也是读者比较关心的问题。

本例中使用了AT29C020 该产品是ATMEL 公司生产的高达256KB 的FLASH ROM，读者也可以使用其它容量不小于256KB 的产品，但在具体电路设计时应注意引脚定义和逻辑关系，并在烧录时选择相应的型号。

UCDOS 软件有一个名字为HZK16 的文件，这就是16x 16 的国标汉字点阵文件，在该文件中按汉字区位码从小到大依次存有国标区位码中的所有汉字，每个汉字占用32 个字节，每个区为94 个汉字，如果显示其属性可以看出其大小为262KB，当写入AT29C020A 时可以将多于256KB 的部分截去，实际上在国标汉字点阵字库中二级汉字排到第87 区，从第88 区开始就是空白，所以将多余部分截去不会影响实际使用。

在实际操作中硬件汉字库的制作和普通的程序存储器的编程没什么不同，注意将HZK16 文件调入时，要使用二进制方式打开，编程并校验成功后即可将存储器取下备用，这样一个硬件汉字库就制作好了。

针对不同的应用读者也可以选用其它大容量甚至是串行数据存储器从而可以扩展硬件汉字库的使用范围。在此也一并向大家推荐：

- 1 容量为512KB 甚至更大的存储器，除了可以存储全部的国标汉字外，还可以存储8x 16 的ASCII 码点阵数据以及汉字语句编码数据，使用起来更加方便。
- 2 串行数据存储器如AT45DB041B 等，特点是占用口线少、体积小和成本低，如果对于速度要求不是很高的场合非常适宜。

三 硬件汉字库使用电路

在本节中笔者使用ATMEL公司的产品89C51 作为主控CPU， 使用控制板VT1335b 来控制320x240 点阵图形液晶显示器，并结合硬件汉字库的使用，将汉字“国”送到液晶屏的（2，0）处显示。

TFT 控制板与单片机的连接信号定义如下：

引脚	符号	功能	备注
1	5V	液晶屏逻辑电源	5V
2	5V	液晶屏逻辑电源	5V
3	D0	数据总线	INOUT 3.3/5V
4	A0	地址线 0	IN 3.3/5V
5	D1	数据总线	INOUT 3.3/5V
6	A1	地址线 1	IN 3.3/5V
7	D2	数据总线	INOUT 3.3/5V
8	CSJ	片选信号，低电平对屏操作有效	IN 3.3/5V
9	D3	数据总线	INOUT 3.3/5V
10	GND	液晶屏逻辑地	0V
11	D4	数据总线	INOUT 3.3/5V
12	GND	液晶屏逻辑地	0V
13	D5	数据总线	INOUT 3.3/5V
14	GND	液晶屏逻辑地	0V
15	D6	数据总线	INOUT 3.3/5V
16	GND	液晶屏逻辑地	0V
17	D7	数据总线	INOUT 3.3/5V
18	RDJ	读操作信号，低电平有效	IN 3.3/5V

武汉东湖新技术开发区珞瑜东路 76 号 电话：027-87680042 138-71489945 137-97051554

www.mcutech.cn

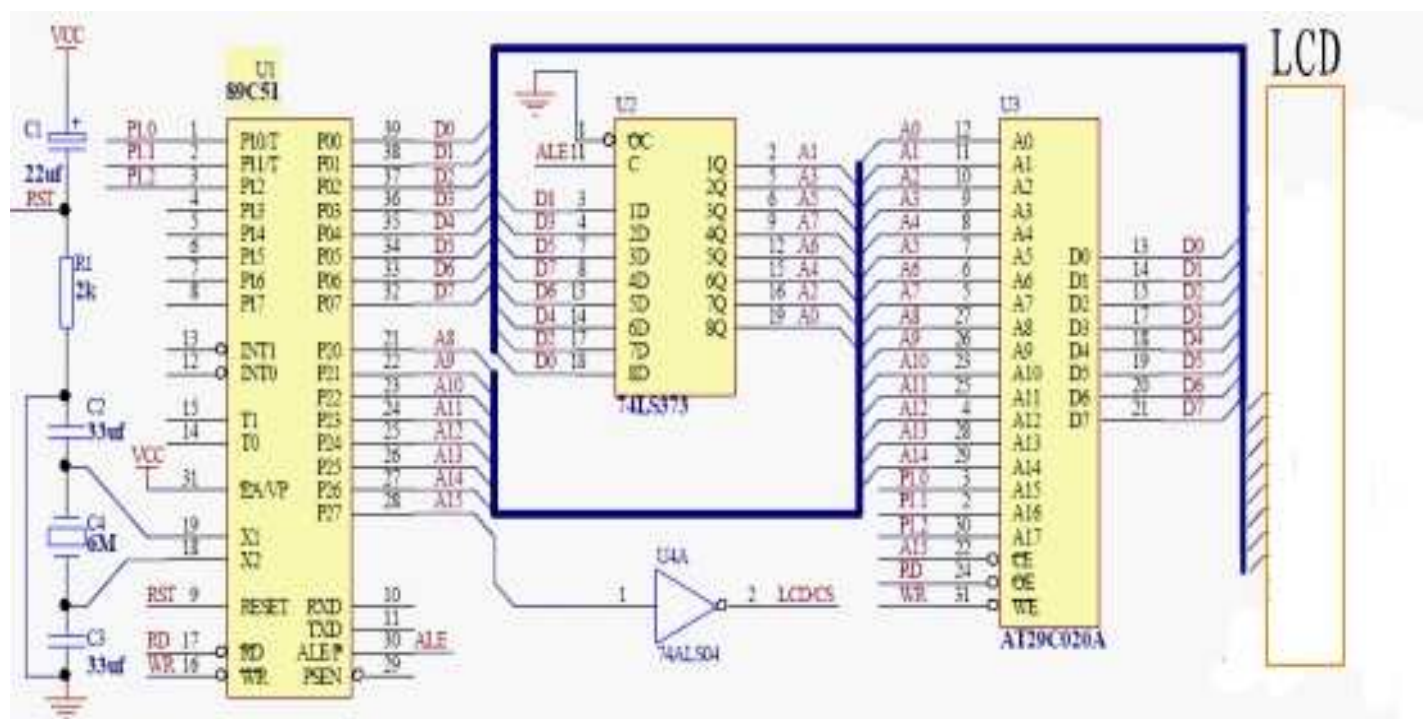
北京中显电子有限公司

www.zxlcd.com

010-82626833,82533328

19	WRJ	写操作信号，低电平有效	IN 3.3/5V
20	NC		

具体电路如图 3 所示



硬件设计说明:

- 1 由于AT29C020 的容量为256KB， 而80C51 系列微控制器只能管理64KB 的数据空间， 所以可以将AT29C020 分成8 页， 每页32KB 占单片机系统数据空间的0000-7FFFH， 剩余32KB 空间为单片机系统的其它存储器和外设。
- 2 A15 反相后作为液晶显示器的片选 LCD/CS。

四 显示原理与软件编制

在第一节中我们已经得出“国”字的区位码为19-5AH， 换算成十进制为25-90， 即“国”字的点阵位于第25 区的第90 个字的位置， 相当于在文件HZK16 中的位置为 $32 * [(25-1) * 94 + (90-1)] = 75040 = 12520H$ 以后的32 个字节为“国”的显示点阵。

在本系统中AT29C020 的A15- A17 分别连接到P1 口的P1.0 P1.1 和P1.2 作为页选择控制， A0-A14 直接连接单片机的A0-A14， 利用MOVX 指令连续取32 个字节送 LCD 的相应位置就能实现正确的汉字显示。

将“国”显示与 (2, 0) 处， 设置VT1335为X方向地址自动加一。 因此， 在每次设置XY地址后， 可以连续写两个字节。 详细程序如下所示。

```
#include <reg51.h>
#include <absacc.h>
#define X_ADDR XBYTE[0x8000]
#define Y_ADDR XBYTE[0x8001]
#define CMD XBYTE[0x8002]
#define DAT XBYTE[0x8003]
unsigned char xdata *flash;//读字库指针
```

武汉东湖新技术开发区珞瑜东路 76 号 电话: 027-87680042 138-71489945 137-97051554

www.mcutech.cn

北京中显电子有限公司 www.zxlcd.com 010-82626833,82533328

```
main()
{
    unsigned long i;
    unsigned char cnt, x, y;
    CMD = 0x04; //x inc, disp on
    //clear panel
    for (y=0; y<240; y++)
    { Y_ADDR = y;
      X_ADDR = 0;
      for (x=0; x<40; x++) DAT = 0;
    }
    i=32 * [ (25-1)* 94 + (90-1) ] //计算“国”字地址
    P1=i/0x8000; //页地址
    flash=i%0x8000; //字库地址
    //读FLASH, 并将数据送到LCD上
    for ( cnt=0; cnt<=15; cnt++ )
    { X_ADDR = 2;
      Y_ADDR = cnt;
      DAT = *flash;
      flash++;
      DAT = *flash;
      flash++;
    }
    while(1);
}
```

五 专用字库芯片

上海集通数码科技有限公司是专业从事汉字库芯片的研发、生产和经营的高科技企业，在中文信息领域具有二十多年的技术经验和积淀，提供多种型号的点阵和矢量字库。硬件汉字库的特点是字库信息完整，包括了国家标准中规定的所有汉字，但硬件成本会略有提高。

参考文献

- [1] 慕声波等. 硬件汉字库设计原理与应用
- [2] 徐卓农, 张永忠. 单片机系统中的汉字显示. 单片机与嵌入式系统应用. 2001(7):31 34
- [3] 李维提, 郭强. 液晶显示应用技术. 第1 版. 北京: 电子工业出版社. 2000.3
- [4] 王士元. C 高级实用程序设计. 第1 版. 北京: 清华大学出版社. 1996.6