# 为什么汉字区位码是94×94不是256×256

计算机只识别由0 1组成的代码，ASCII码是英文信息处理的标准编码，汉字信息处理也必须有一个统一的标准编码。 汉字交换码（区位码）主要用于汉字信息交换，我国国家标准局于1981年5月颁布了《信息交换用汉字编码字符集——基本集》，代号为GB2312-80，共对6763个汉字和682个图形字符进行了编码，其编码原则为：汉字用两个字节表示（ASCII码用一个字节），原则上，两个字节可以表示 256×256=65536 种不同的符号，作为汉字编码表示的基础是可行的。但考虑到汉字编码与其它国际通用编码，如ASCII 西文字符编码的关系，我国国家标准局采用了加以修正的两字节汉字编码方案，只用了两个字节的低7位。这个方案可以容纳 128×128=16384 种不同的汉字，但为了与标准ASCII码兼容，每个字节中都不能再用32个控制功能码（什么是控制功能码？见下文“什么是ASCII功能码”）和码值为32（二进制数为：100000）的空格以及127（二进制数为：1111111）的操作码。所以每个字节只能有94个编码。这样，双七位实际能够表示的字数是：94×94=8836个。

## 什么是ASCII功能码：

ASCII 字符集，最基本的包含了 128 个字符。其中前 32 个， 0-31 ，即 0x00-0x1F ，都是不可见字符。这些字符，就叫做控制字符。

这些字符没法打印出来，但是每个字符，都对应着一个特殊的控制功能的字符，简称功能字符或功能码 Function Code 。

此外，由于 ASCII 中的 127 对应的是 Delete ，也是不可见的，所以也可以归为 Function Code 。

关于每个控制字符的控制功能缩写，参见下表：

**表格  1    ASCII 中的控制字符**

|  |  |  |  |
| --- | --- | --- | --- |
| **十****进制** | **十六****进制** | **转义****字符 \*** | **说明** |
| 0 | 00 | /0 | Null character( 空字符 ) |
| 1 | 01 |   | Start of Header( 标题开始 ) |
| 2 | 02 |   | Start of Text( 正文开始 ) |
| 3 | 03 |   | End of Text( 正文结束 ) |
| 4 | 04 |   | End of Transmission( 传输结束 ) |
| 5 | 05 |   | Enquiry( 请求 ) |
| 6 | 06 |   | Acknowledgment( 收到通知 / 响应 ) |
| 7 | 07 | /a | Bell ( 响铃 ) |
| 8 | 08 | /b | Backspace( 退格 ) |
| 9 | 09 | /t | Horizontal Tab( 水平制表符 ) |
| 10 | 0A | /n | Line feed( 换行键 ) |
| 11 | 0B | /v | Vertical Tab( 垂直制表符 ) |
| 12 | 0C | /f | Form feed( 换页键 ) |
| 13 | 0D | /r | Carriage return( 回车键 ) |
| 14 | 0E |   | Shift Out( 不用切换 ) |
| 15 | 0F |   | Shift In( 启用切换 ) |
| 16 | 10 |   | Data Link Escape( 数据链路转义 ) |
| 17 | 11 |   | Device Control 1( 设备控制 1) /XON(Transmit On) |
| 18 | 12 |   | Device Control 2( 设备控制 2) |
| 19 | 13 |   | Device Control 3( 设备控制 3) /XOFF(Transmit Off) |
| 20 | 14 |   | Device Control 4( 设备控制 4) |
| 21 | 15 |   | Negative Acknowledgement( 拒绝接收 / 无响应 ) |
| 22 | 16 |   | Synchronous Idle( 同步空闲 ) |
| 23 | 17 |   | End of Trans the Block( 传输块结束 ) |
| 24 | 18 |   | Cancel( 取消 ) |
| 25 | 19 |   | End of Medium( 已到介质末端 / 介质存储已满 ) |
| 26 | 1A |   | Substitute( 替补 / 替换 ) |
| 27 | 1B | /e | Escape( 溢出 / 逃离 / 取消 ) |
| 28 | 1C |   | File Separator( 文件分割符 ) |
| 29 | 1D |   | Group Separator( 分组符 ) |
| 30 | 1E |   | Record Separator( 记录分隔符 ) |
| 31 | 1F |   | Unit Separator( 单元分隔符 ) |
| 32 | 20 |   | White space |
| 127 | 7F |   | Delete( 删除 ) |