

电脑快速入门丛书

DIANNAOKUAISURUMENCONGSHU

实用汉字输入 打印和显示方法

齐仲修 柯斌 郭哲 编著

简明、准确、实用
简明、准确、实用

给初学者的书
给自学者的书



青岛出版社

25278

电脑快速入门丛书

实用汉字输入、打印和显示方法

齐仲修 柯斌 郭哲 编著

青岛出版社

鲁新登字 08 号

责任编辑 樊建修
封面设计 李伯书

实用汉字输入、打印和显示方法

齐仲修 柯斌 郭哲 编著

*

青岛出版社出版

(青岛市徐州路77号)

邮政编码: 266071

新华书店北京发行所发行

胶州市装潢印刷厂印刷

*

1995年6月第1版 1995年6月第1次印刷

32开(850×1168毫米) 4.75印张 115千字

印数 1—20110

ISBN 7-5436-1283-6/TP·99

定价:5.40元

目 录

第一章 汉字处理技术简介	1
第一节 汉字处理技术简介	3
国家标准汉字字符集	3
汉字点阵和字模	4
汉字的输入	6
汉字的各种编码	7
汉字操作系统	8
第二节 2.13H 汉字系统	9
系统安装	9
2.13H 的启动	12
2.13H 的汉字输入功能	16
2.13H 的打印功能	20
2.13H 的特殊显示功能	25
第二章 常用汉字输入方法	32
第一节 区位码输入法	32
第二节 简化拼音码输入法	33
第三节 双拼双音输入法	34
什么是双拼	34
什么是双音	34
输入多字词汇	37
自定义词组	37
联想输入方式	39
二级简码汉字	40
第四节 五笔字型输入法	44

熟悉字根	45
汉字拆分练习	51
简码和词组	56
常见非基本字根的拆分	61
第三章 自然码输入法	65
第一节 自然码的进入和退出	65
自然码的进入	65
自然码的退出	66
第二节 怎样输入单字	67
简码字输入	67
拼音输入	67
拼音码加部首码输入	69
第三节 怎样输入词组	71
简码词输入	71
用拼音方式输入双字词	71
三字词的输入	72
四字词的输入	72
四字以上词的输入	73
第四节 怎样增加和使用自造词组	73
如何增加自造词组	74
删除词组	74
保存词组	75
装入自造词组	75
第五节 特殊输入功能	76
联想输入方式	76
南方音方式	77
选择不同的提示方式	77
输入制表符	78

输入中文数字、年月日等	78
输入中文标点	78
输入罗马数字、俄文、日文假名等	79
第四章 五十字元汉字输入法	80
第一节 字元	80
第二节 形声编码	84
基本规则	84
引申规则	87
多音字和不会读音的字的输入方法	89
可能引起取元误解和费解的字	89
高频简码字	92
词语编码	92
第三节 声形编码	94
第四节 全形码	95
双字元字和三字元字的编码	95
四字元以上的字的编码	96
四组例外字的编码	97
交叉结构的字的编码	98
词语编码	99
第五节 音形编码	101
字元字和交叉结构的字的编码	101
双字元字的编码	101
三个字元以上的字的编码	101
词语编码	103
第六节 形音编码	104
【附录一】 2.13H 字型样张	105
【附录二】 区位码—五笔字型码索引速查表	107

第一章 汉字处理技术简介

世界上的文字可以分成拼音文字和表意文字两大体系。

在拼音类文字中,数以万计的单词和词组是由数量不多的若干个字母排列组合构成的。例如英文就是由 26 个字母组成的。由于构成文字的基本字元很少,信息的存储、输入、处理、输出都很好解决,比较容易用计算机处理。

表意类文字是人类历史上最早出现的一种文字,例如古埃及文字、古玛雅文字及汉字……,汉字是唯一流传并巩固下来的一种表意文字。表意文字用各种独立的符号来构成词和句,与语音没有直接联系,随着社会的发展,语言词汇不断丰富,汉字所包含的符号(单字)的数量与日俱增。见下表

我国历代字书收字数量

时代	字书名	收字数量
西汉	仓颉篇	3300
东汉	说文解字	9350
晋	字林	12800
唐	广韵	26100
宋	类篇	31300
明	正字通	33400
清	康熙字典	49174
1986 年	汉语大字典	约 60000

汉字总字数虽然很多,但人们经常使用的汉字字数是 5000~

7000,而且每个字的使用频度相差很大。《毛泽东选集》四卷共使用 2975 种汉字。1977 年,新华社、北京新华印刷厂曾组织 19 个单位,历时两年,对 86 本书、104 种期刊、7075 篇文章、总计 2100 多万字进行统计,这么多书刊文章共用汉字 6347 种。在这 6347 种汉字中,“的”一个字就用了 83 万多次,占 4%强。而频度低的两千个字总共使用一万次,合起来才占 0.05%,同“的”字相比频度相差 16 万倍。

汉字种类繁多、构造复杂的特点曾给计算机处理带来很大困难。直至 70 年代,国内外专家还都普遍认为:“为了适应现代科学技术的发展,汉字改革要走世界文字共同的拼音方向,传统的方块汉字最终将被淘汰。”但是 80 年代以来,由于大规模集成电路存储器的出现,汉字处理技术飞跃发展,从根本上改变了这种状况。在电脑上实现汉字信息处理,不仅十分方便,而且同西文处理相比,反而具有某些优势。

1. 目前,一台较低档的 286 机都配置 1M 内存和 40M 硬盘,汉字字库多占用 2~3M 硬盘空间已不成为技术问题。而英文单词通常要比中文长,例如:“信息”一词,汉字占 4 个字节,而英文 information 占 11 个字节;“语言”一词,汉字占 4 个字节,而英文 language 占 8 个字节……。同一篇文章,汉字文本的长度通常只是其英文文本的 $1/2 \sim 2/3$,著名小说《简·爱》,其英文版约 120 万个英文字符,而中文版只有 40 万个汉字(相当于 80 万个英文字符)。综合比较,汉字信息存储可以更有效地节省电脑的存储空间。

2. 汉字输入技术的发展大大提高了汉字输入速度。英文单词的平均长度约 5 个字母,加上词与词之间的空格,输入一个英文单词平均需击键 6 次。近几年推出的一些汉字输入方案,如双拼双音、五笔字型、自然码、声数码等,输入一个单字最多需击键四次,而且这些方案都提供了方便实用的联想及词组功能,可以

快速输入词组，因此稍加培训即可达到每分钟 100 字的水平。实践证明，汉字输入并不比英文输入慢。

3. 全世界总人口的 36% 使用汉字，除了中国，日本、南朝鲜、东南亚各国也都使用汉字，用汉字记载的光辉灿烂的中国古代文化是全人类的宝贵遗产。泛泛地讲汉字比拼音文字难学的说法是不能成立的——汉字虽多，但只要认识 700 左右就能写日常短文。而英文单词总数约 160000 个，认识 700 个单词连菜单都看不懂。以“羊”字为例，汉字认识一个“羊”字，就可以联想到羊族中的其他词，如公羊、母羊、羊毛、羊肉等。而英文中羊的总称是 sheep、公羊是 ram、母羊是 ewe、山羊是 goat、羊毛是 wool、羊肉是 mutton，毫无联系。

近年来，国内外许多专家指出：“汉字是一种科学的、先进的文字，21 世纪将是汉字汉语发挥威力的时代。”古老的汉字在新技术革命的浪潮中不仅不会被淘汰，反而焕发出新的生机，成为信息储存与传递的有力工具。

第一节 汉字处理技术简介

一、国家标准汉字字符集

根据使用频度可以把汉字划分为高频字(约 1000 个)、常用字(约 3000 个)、次常用字(约 4000 个)、罕见字(约 8000 个)和死字(约 45000 个)。

为了适应汉字信息处理的需要，我国 1981 年颁布了国家标准《信息交换用汉字编码字符集——基本集》(GB2312-80)，俗称“国标字符集”或“GB2312-80 字符集”。目前我国研制的各种中文处理系统均以此为标准。这个字符集共收入汉字 6763 个，

收集了高频字、常用字及绝大部分次常用字，按频度分成两级。第一级有 3755 个汉字，按汉语拼音顺序排列，同音字再以笔划排序(笔划顺序是：横、竖、撇、点、折)；第二级有 3008 个汉字，按偏旁部首顺序排列。

国标字符集还收入各种字母和符号 682 个，其中：

一般符号 202 个，包括标点符号、运算符号、单位符号和制表符号等

序号 60 个：1. ~ 20. , (1) ~ (20), ① ~ ⑩, (一) ~ (十)

数字 22 个：0 ~ 9, I ~ XI

拉丁字母 52 个：A ~ Z, a ~ z

日文假名 169 个：平假名 83 个，片假名 86 个

希腊字母 48 个：A ~ Ω, a ~ ω

俄文字母 66 个：A ~ Я, a ~ я

汉语拼音符号 26 个

汉语注音符号 26 个

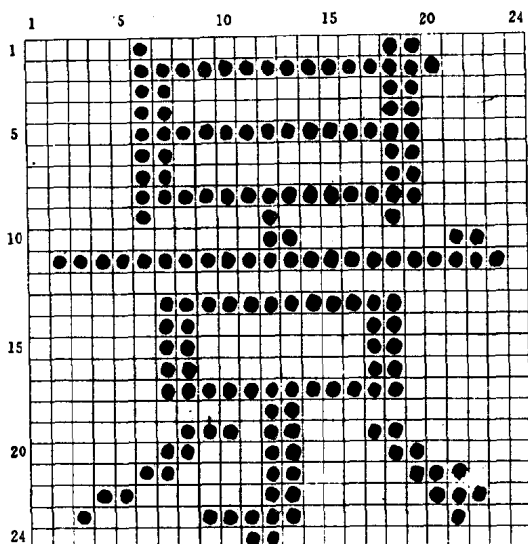
目前正在拟定的国家标准字符集在基本集的基础上，增加了第二辅助集 7237 个，第四辅助集 7039 个，共收汉字 21039 个。第一、三、五辅助集分别是与基本集、第二辅助集、第四辅助集对应的繁体字。

二、汉字点阵和字模

可以把汉字看成是一个二维图形。为了便于描述和处理，通常是把一个汉字离散成网点，每个点用一个二进制位表示，从而组成该汉字的一个点阵，如下图所示。描述一个汉字点阵信息的一串字节称为该汉字的字模。

点阵规模用“横向点数×纵向点数”表示。从印字质量看，当然是点数越大越好，但这会使设备成本增加。目前，用于屏幕显示的点阵规模是 16×16，用于点阵打印机的点阵规模是 24×24

或 48×48 ，而排版印刷则采用 128×128 或更大规模的点阵。



一个汉字的 24×24 点阵图案

汉字字模库(简称字库)用来存储一个个汉字的点阵信息。一个 16×16 点阵汉字的字模占 32 个字节，一个 24×24 点阵汉字的字模占 72 个字节，显然，为了实现近 7000 个汉字的显示和打印，这个字库需占用很大的存储空间。 16×16 点阵字库约占 238K 字节， 24×24 点阵字库约占 536K 字节。

字库可以存放在系统内存(RAM)中，由于 RAM 中存储的信息断电后立即消失，因此每次开机时都需要从磁盘上把字库装入系统内存，这种字库称为软字库。软字库实现起来比较简单，不

需要改动机器的硬件，应用很广泛。其缺点是要占用一部分系统内存。

另一种做法是把汉字字库与系统内存分开，用 EPROM 或 Mask-ROM 芯片制成汉字字模库卡（俗称汉卡），把它插入机器的扩充槽中，作为机器的一个扩充 ROM。这种方法不占用系统内存，近年来应用得很广泛。

为了减少字库空间，也有人在信息存储方法上做文章并取得了很好的效果，如矢量式字库、霍夫曼树形压缩字库、字元式压缩字库等。

三、汉字的输入

如何输入汉字？是一个汉字处理系统必须首先解决的问题。许多专家从事这方面的研究，提出了各种各样的输入方法，这些方法归纳起来可分为三类——编码输入，借助专用“汉字键盘”的整字输入，利用语音识别装置或文字识别装置的自然语言输入。

编码输入是用字母数字串来表示汉字，因而可以直接使用常用的西文键盘输入汉字，是汉字输入方法的主流。目前已研制出的编码方案大约有 400 多种，其中 40 多种已在机器上实现，这些编码大体上可分成四类：

(1) 用四位数字表示一个汉字或符号，例如区位码、电报码。这种编码的特点是无重码，编码对应的汉字是唯一的。

(2) 按照汉语拼音对汉字进行编码，例如拼音码、双拼双音码、自然码。

(3) 按照汉字的字形结构进行编码，如首尾码、五笔字形。

(4) 把汉字的发音和字形结合起来编码，例如音形码、长城 ABC 码。

这些方案各有千秋，本书仅介绍最常用的几种。

四、汉字的各种编码

在机器中，西文字符只有一种编码，目前最常用的是 ASCII 码。在 ASCII 码中，一个西文字符用 7 位二进制数表示，如字母 A 的 ASCII 码是 1000001，占一个字节(第 8 位是 0)。而在汉字处理系统中，为了便于在系统的各个环节表示和处理汉字，要使用不同的汉字编码，主要有以下几种：

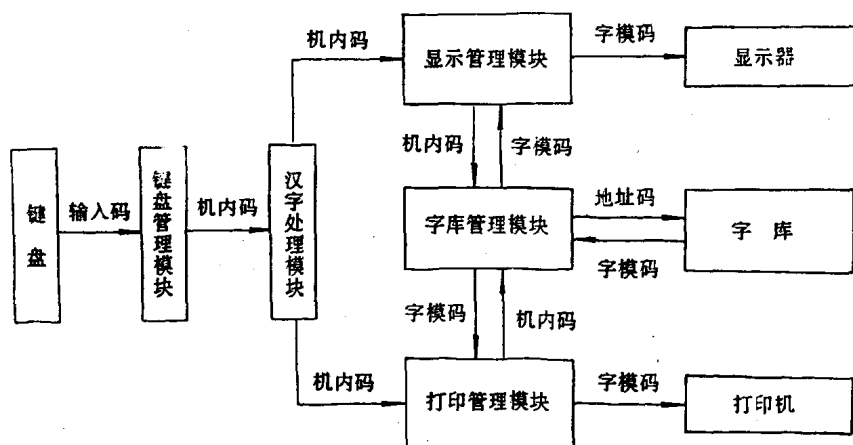
1. 输入码：是用户从键盘输入汉字时使用的编码。同一个汉字在不同的输入方法中，输入码大不相同，例如：“啊”字在区位码中的输入码是“1601”，在简化拼音中的输入码是“a”，在五笔字型中的输入码是“kbsk”。

2. 机内码：是机器内部表示汉字的代码。一个汉字的机内码占两个字节，为了避免与西文 ASCII 码混淆，这两个字节的最高位均为“1”。例如，汉字“啊”的机内码为 10110000, 10100001。

3. 字模码：是每个汉字的点阵信息，简称字模。一个 16×16 点阵汉字的字模占 32 个字节，一个 24×24 点阵汉字的字模占 72 个字节。

4. 地址码：是每个汉字的字模(点阵信息)在字库中的起始字节地址。

汉字的处理过程就是各种汉字代码互相转换的过程。如下图所示：输入时，用户从键盘把输入码送入机器，键盘管理模块把它转换成统一的机内码。输出时，先通过字库管理模块把机内码变成地址码，然后按地址码从字库中取出字模码，再通过显示管理模块送屏幕显示，或通过打印管理模块送打印机打印。



汉字编码传送示意

五、汉字操作系统

DOS 磁盘操作系统是“舶来品”，不能处理汉字，这严重地阻碍了电脑在中国的使用和普及。1983 年 12 月，电子工业部六所推出我国第一个汉字操作系统 CCDOS(Chinese character Disk Operating System)，开创了汉字处理技术的新局面。

汉字操作系统是在西文 DOS 的基础上，对其进行扩充和改造而形成的，它主要是扩充了 DOS 基本输入输出系统(BIOS)的键盘管理模块、显示管理模块、打印管理模块、字库及字库管理模块，使之具有汉字输入、汉字显示、汉字打印以及汉字文件的存贮、传送等能力。

汉字操作系统只是给西文 DOS 增加了汉字处理功能，在使用上与西文 DOS 完全兼容，可以直接使用西文 DOS 的各种命令

和功能。对用户来说，汉字和西文一样，也是一种字符，在各种程序语言和应用程序中都可以象使用西文一样地去使用汉字。但要注意的是：一个汉字占两个字节，相当于两个西文字符。例如，DOS 规定文件的基本名可取 8 个字符，扩展名可取 3 个字符，若用汉字作文件名，则基本名最多可取 4 个汉字，扩展名只能用 1 个汉字。

近年来，汉字处理技术飞跃发展，推出了许多性能优异的汉字操作系统，如：电子工业部六所研制的 CCDOS4.0，吴晓军先生研制的 2.13H，金山公司研制的 SPDOS，王永民先生研制的 WMDOS，联想公司研制的 LXDOS，希望公司研制的 UCDOS...，对用户来说，这些汉字操作系统的使用方法基本类似，本章介绍流行面较广的 2.13H 汉字操作系统。

第二节 2.13H 汉字系统

2.13H 汉字系统适用于 CPU 为 8088、80286、80386 及 80486 的各种电脑，如 IBM 系列、COMPAQ 系列、AST 系列、SUN 系列、ALR 系列、VST 系列、DEC 系列、SUPER 系列……，要求内存 $\geq 512K$ 且配有硬盘。支持各种常用的显示器和打印机，可以在 2.0 以上的各种 DOS 版本下工作。

一、系统安装

2.13H 的基本系统(包括宋、仿宋、黑、楷四种字体的 24 点阵字库)约占 2.5 兆字节，繁体字库约占 760K 字节，高点阵字库约占 5 兆字节。在安装前要先检查硬盘上是否有空间，如果空间不够，则需先删除一些无用文件后再安装。

2.13H 只能安装到 C 盘上，在安装过程中它自动在 C 盘根

目录中建立一个名为 213 的子目录, 并把大部分文件拷贝到 213 子目录中, 只有 AUTOEXEC. BAT、CONFIG. SYS、ANSI. SYS 和安装批文件 ZK. BAT(或 ZKB. BAT)存在根目录下。

2. 13H 系统的安装步骤如下: 把 7 号盘插入 A 驱动器, 然后敲入以下命令(如果需要从 B 驱动器安装, 敲入命令 ZKB $\swarrow$)

ZK $\swarrow$

屏幕上就会提示你依次插入 0 号~6 号盘。在安装过程中, 要求用户选择显示器的类型, 屏幕显示如下

1—CGA(640×200)

2—EGA(640×350)

3—VGA(640×480)

4—CGE400(640×400)

5—HERCULES(720×350 单显)

请根据显示器的类型, 敲入 1~5 中的一个数字。

最后一张盘装入后, 要选择打印机

1—P1351, P1350, P350 彩色

2—M2024, M1724, 1900, 1924

3—TH3070, KC3070, 紫金 3070, 紫金 3080,

TH3070SL, TH3070R1/R2

4—AR2463

5—LQ1000(K), LQ1500(K), LQ1600K, VP8624,

LQ2500K 彩色, M1724L, NECP7

6—OKI8320(C), OKI8330C, OKI5320C

7—M1570 彩色

8—NEC3824, PR201(P5)

9—NM9400, NEC9400

0—AR3240

请根据打印机的类型, 敲入 0~9 中的一个数字。

【说明】 系统安装完毕后，如果要重新选择打印机的型号，可敲入以下命令：

```
C>CD 213 ✓
```

```
C>PRTA □ ✓    (“□”表示空格键)
```

屏幕上显示出以上打印机选择菜单，这时，再根据打印机的类型，敲入一个数字即可。

(1) 安装繁体字库

如果要安装繁体字库，需要再执行以下操作：

```
C>CD \ 213 ✓
```

把繁体打印字库的 1 号盘插入 A 驱动器，

```
C>RESTORE A: C: ✓
```

根据提示再插入繁体打印字库的 2 号盘。

```
C>MD FT ✓
```

```
C>CD FT ✓
```

把繁体显示字库盘插入 A 驱动器，再依次执行以下命令

```
C>COPY A: *.* ✓
```

```
C>COPY \ 213\MENUHH.COM ✓
```

```
C>COPY \ 213\FILE *.* ✓
```

```
C>CD \ ✓
```

```
C>COPY A:FT.BAT ✓
```

(2) 安装高点阵字库

高点阵字库包含宋体、仿宋体、黑体、楷体、图形符号五个 40×40 点阵字库，共 14 张盘，用 RESTORE 命令来安装，如下：

```
C>CD \ 213 ✓
```

```
C>RESTORE A: C: ✓
```

根据提示依次插入字库盘。