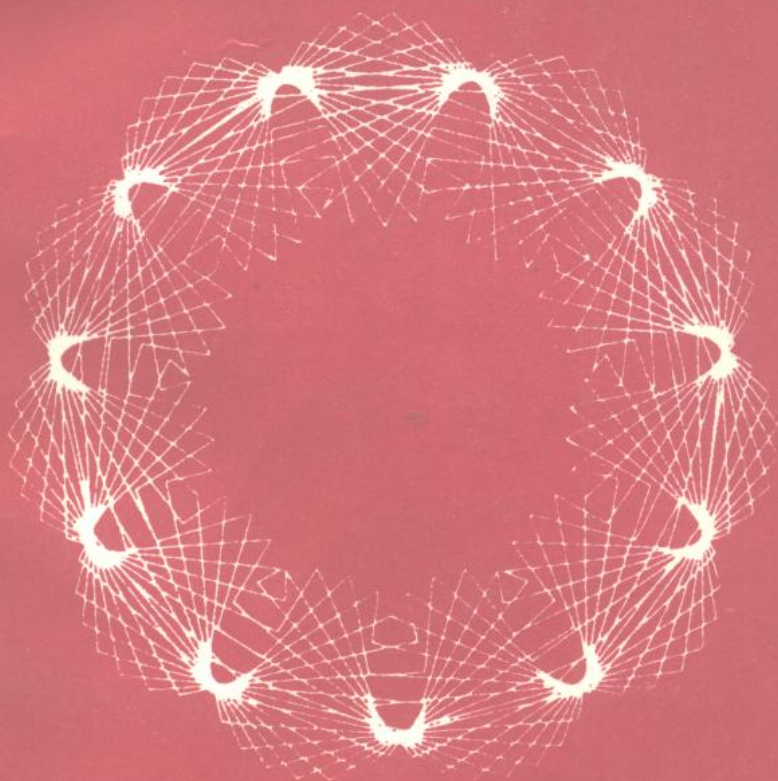


● 计算机应用基础系列丛书

计算机常用汉字

输入方法

王盼卿 王礼成 董淑英 编



91.14
2/1

电子工业出版社

计算机应用基础系列丛书

计算机常用汉字输入方法

王盼卿 王礼成 董淑英 编

电子工业出版社
9023703

(京)新登字 055 号

内容提要

本书介绍了十余种常用的计算机汉字输入方法,及计算机键盘输入指法练习。

主要内容有:五笔字型、自然码、大众码、国标码、区位码、电报码、汉语拼音、拼音联想、首尾及快速等十余种汉字输入法,以及键盘指法训练等。

本书适合于计算机应用及操作、使用人员阅读,也可作为计算机操作培训班教材。

JS86/05

计算机常用汉字输入方法

王盼卿 王礼成 董淑英 编

责任编辑 王昌铭

电子工业出版社出版(北京市万寿路)

电子工业出版社发行 各地新华书店经售

机电部情报所印刷

开本: 850×1168 毫米 1/32 印张: 5.25 字数: 132 千字

1992 年 10 月第 1 版 1993 年 3 月第 2 次印刷

印数: 7100—27200 册 定价: 3.80 元

ISBN7-5053-1923-X/TP·459

前 言

随着计算机技术的迅速发展,计算机的应用范围日益广泛。为适应计算机应用工作的普及和发展的需要,我们编写了本丛书。本丛书共十本,涉及到计算机基础知识的各个方面,内容丰富、实用性很强,是全面学习和掌握计算机应用难得的普及性读物,其内容包括:《操作系统及应用》、《微型机常用通用软件》、《计算机常用汉字输入方法》、《计算机硬件基础》、《计算机软件基础》、《C语言应用基础》、《ORACLE数据库应用基础》、《dBASE数据库应用基础》、《计算机网络应用基础》、《微型机常见故障维修实例》。

在我国汉字输入技术是计算机应用工作的基础,不掌握汉字输入,计算机的普及在我国就不能真正实现。在编写本书时,我们本着通俗易懂、深入浅出的原则,力求简明、实用。书中汇编了国内常用的十余种汉字键盘输入方法,它可作为各类微机培训班的教材,亦可作为各层次人员的自学和参考手册。

本书第一、四、五章由王盼卿编写,第二章由王礼成编写,第三章由董淑英编写。全书由陈致明、王盼卿统编,由陈致明副教授、木林林、王大刚主审。

本书在编写过程中,军械工程学院自动化系胡晓初主任、田秋成副主任给予了极大关心,总后军械部周绍钧处长以及木林林、王大刚同志给予了大力的支持和指导,在此一并致谢。

由于编者水平有限,误漏之处在所难免,恳请批评指正。

编者

1992年4月

目录

第一章 汉字输入技术概述	(1)
§ 1.1 计算机与汉字输入	(1)
§ 1.2 汉字编码输入技术的发展	(2)
§ 1.3 汉字输入方法选择要点	(3)
§ 1.4 常见的几种汉字输入方法之比较	(4)
第二章 五笔字型汉字输入法	(9)
§ 2.1 五笔字型概述	(9)
§ 2.2 五笔字型字根	(16)
§ 2.3 五笔字型的输入方法	(30)
§ 2.4 汉字输入举例	(47)
§ 2.5 五笔划与五笔桥输入法	(49)
§ 2.6 造词与造字软件使用说明	(55)
第三章 自然码汉字输入法	(58)
§ 3.1 自然码概述	(58)
§ 3.2 自然码的汉字编码规则	(60)
§ 3.3 自然码汉字输入示例	(73)
§ 3.4 自然码的特殊功能	(96)
第四章 其他常用的汉字输入方法	(103)
§ 4.1 区位码输入法	(103)
§ 4.2 国标码输入法	(104)
§ 4.3 电报码输入法	(104)
§ 4.4 汉语拼音输入法	(104)
§ 4.5 拼音联想输入法	(109)
§ 4.6 首尾码输入法	(110)

§ 4.7	快速输入法	(112)
§ 4.8	大众码输入法	(113)
第五章	键盘指法训练	(117)
§ 5.1	指法训练要点	(117)
§ 5.2	键盘指法练习	(119)
附录一	汉字区位码表(1—9区)各种符号	(131)
附录二	五笔字型汉字拆分示例	(135)
附录三	自然码简词表	(149)
附录四	自然码键盘对照表	(151)
附录五	自然码双拼方案与全拼方案键盘对照表	(152)
附录六	自然码的双拼声母	(152)
附录七	自然码的双拼韵母	(152)
附录八	自然码中文标点方式下的中文标点键盘对照表	(152)
附录九	自然码“义部部首”及“非部件字形码”键盘表	(153)
附录十	自然码需要特别记忆的字形码表	(153)
附录十一	自然码汉字输入系统数据管理程序的使用	(153)
附录十二	自然码汉字输入系统自定义字词管理程序 使用说明	(156)

第一章 汉字输入技术概述

§ 1.1 计算机与汉字输入

汉字是世界上最古老的文字之一,是中国人民宝贵的文化遗产;汉字是世界上使用人口最多的文字,也是联合国的工作语言之一。

自本世纪六十年代以来,电子计算机以其神奇的功能给人类社会带来了空前的社会文明。目前,在一些发达国家,计算机已广泛应用于政治、军事、经济、文化、科技等各个领域。无论是在科学计算等数值计算方面,还是在办公自动化等非数值信息处理方面,都发挥着越来越大的作用。在我国,电子计算机的应用从六十年代就开始了,直到进入八十年代以后,才得到广泛的应用。但是,要使电子计算机技术在我国真正的普及和推广,进入普通的办公室和家庭,就必须解决电子计算机的中文信息处理问题,并在此基础上,提高计算机对汉字中文信息的处理能力。计算机中文信息处理的关键是汉字信息的输入问题。汉字进入计算机是中文信息处理的基础,没有这个基础,其他就都谈不上了。

为了使古老的汉字能够迅速、高效地适应信息技术的发展,使汉字能进入计算机并可参与信息处理,使我国的计算机应用走向普及,七十年代以来,具有聪明才智的炎黄子孙,对计算机的汉字输入技术进行了深入的研究与开发。其输入技术大致可分为如下三方面:汉字编码输入技术、汉字识别输入技术、汉字语音输入技术等。仅就汉字编码输入技术而言,到八十年代中期,据有关资料

统计,已有汉字输入编码方案和设想五百多种。这就为汉字适应电子计算机的环境,进行计算机处理创立了有效的手段。

§ 1.2 汉字编码输入技术的发展

随着计算机技术的普及和推广,经过数年的探索和实践,有关人士已逐渐认识到,要想实现真正的办公自动化,仅靠配备专业操作人员将永远无法真正提高办公效率,也无法实现中文信息处理的现代化,而必须让人们放下手中的笔和纸,将汉字直接输入计算机,使汉字输入工作基本上一次性完成。这也是汉字编码输入技术的目的和发展方向。

1.2.1 汉字编码方案的发展

汉字编码方案的发展,从总体上来说,是从无规律到有规律,从难记到易记,从低效率到高效率的。汉字编码的分类如下:

1. 无理序号编码

纯无理序号编码在我国没有出现过。这种汉字编码最短,但记忆难度极大。此种方法若经长时间的训练,熟练掌握后,其输入速度从理论上说是最快的。

2. 分类序号编码

典型的分类序号编码包括:国标码、区位码、电报码。

3. 有理编码法编码

主要分四类:拼音类、音形结合类、笔面拼形类和字元拼形类。

包括:汉语拼音、拼音联想、首尾码、五笔字型、自然码、大众码等数百种。

1.2.2 汉字输入的时代划分

第一代:以单字输入为主。

早期发明的几百种输入法大多以此为基础。

第二代:以词组输入方式为主。

以字为基础,词为主导,音形结合,字词一体,高频先见智能处理的输入方式属于第二代。象自然码基本上属于第二代产品。目前,常用的《五笔字型》、《大众码》等汉字输入方法是以单字为主,但兼有部份词的功能,是介于第一代与第二代之间的产品。

随着计算机技术的发展,汉字编码输入效率已由过去的依赖编码转向依赖系统,依赖智能化的软件设计,同样的编码在不同的软件支持下将会有明显的差异。

第三代:以语句智能化语音输入为主,必须通过第五代计算机去完成。

§ 1.3 汉字输入方法选择要点

选用哪一种方法输入汉字,使用者可根据自己的实际情况择优选用。选择时应注意如下要点:

1. 编码规则要简单明确,易学易记,便于熟练掌握。
 2. 编码容量要大,至少应覆盖 GB2312-80 基本集中规定的全部汉字(6763 个),并易于扩充。在必要情况下,应能覆盖辅助汉字集中所包括的汉字。
 3. 重码率要低,并有简便的区分重码的方法。
 4. 编码要有唯一性。同一个汉字只能产生一个确定的编码。
 5. 尽可能使用标准字符键盘作为汉字输入的设备,键位的分布合理。
 6. 汉字输入的平均击键次数要少,尽量缩短编码长度,以提高编码效率。
 7. 输入速度要高。
- 总之,要选择易学、易记、易用、输入速度高的输入方法。

§ 1.4 常见的几种汉字输入法之比较

目前,各种汉字输入法有数百种之多,但真正商品化的仅几十种;在计算机(特别是微型计算机)上应用面较广的几种汉字操作系统中,主要配备了:国标码、区位码、电报码、汉语拼音、拼音联想、首尾及快速、大众码、五笔字型、自然码等汉字输入方法。这些输入方法是本书主要介绍的内容。

为便于大家选择适合自己实际情况的一两种输入法,下面对这几种输入法进行简单的介绍和比较。

1.4.1 区位码

区位码输入法是按国标 GB1988-84 和国标 GB2311-84 中对各种汉字和符号的编码规定,对国标 GB2312-80 中 7445 个汉字和符号,按行、列进行划分,共划分为 94 个区(行)和 94 个位(列)。每个汉字或符号都对应于一个区号和位号。输入时,只要输入对应的区号和位号四位数字,即可得到对应的汉字或符号。

该方法的优点是:只要有一本区位码表,不需任何训练,即可输入汉字和各种符号。码长为 4 键/字。

主要缺点是:即使进行专门训练,也不可能完全记住区位码表,而查表则严重影响输入速度。

1.4.2 国标码

国标码是区位码的另一种表示方法,国标 GB2312-80 中每个汉字和符号用四个数字或字母(A、B、C、D、E、F)来表示,即用四位十六进制数来表示。输入时,对照国标码表,查出对应的编码,并依次输入,则相应的汉字或符号就输入计算机了。

例如:“摧”字的编码为 345D,依次按下数字键“345”和字母键

“D”，“摧”字就输入计算机了。

用国标码输入汉字的优缺点和区位码完全相同，只是不如区位码使用面广。

1.4.3 电报码

电报码是按电报明码编码法，用四位数字对应一个汉字。输入时，先查到汉字对应的编码，依次输入相应的数字即可。

例如：“多”字的电报码为 1122，依次按下数字键“1122”，汉字“多”就输入计算机了。

用电报码输入汉字的优缺点和区位码基本相同。

1.4.4 汉语拼音

汉语拼音方法是微型计算机上早期实现的一种汉字输入方案。现在，几乎所有的通用汉字操作系统和专用操作系统中均配有该方法。该方法是用输入汉字对应的汉语拼音字母来选择和输入汉字的。据统计，中文汉字约有 50000 个，国标码中规定的 1、2 级字库汉字有 6763 个，只有不到 500 种不同的拼音；显然，拼音相同的汉字较多，输入时只能从同音字中选字。

汉语拼音法的主要优点是：学习时间短，对于会汉语拼音的人来说，不需专门训练，即可向计算机中输入汉字。

主要缺点是：输入速度慢，不认识的汉字无法输入且汉字库中的 682 个符号无法输入。

1.4.5 拼音联想输入法

拼音联想输入法是基于拼音输入法又带有联想功能的一种汉字和符号的输入方法。当采用此法输入时，拼音联想系统自动调整汉字及联想字的位置，保证最后输入的汉字自动调换到同音字的最前面。从而大大减少选字的时间。

汉字联想输入法的主要优点是：速度比较快，学习掌握很容易，不需要专门训练，并可输入汉字库中的常用符号。对于输入汉字的同时又需要输入各种符号的操作尤为方便。

主要缺点：基本同于汉语拼音输入法。

1.4.6 首尾码及快速输入法

首尾码是用 52 个字根来概括组成汉字的基本笔画。52 个字根用键盘上的 26 个英文字母来对应。输入汉字时，先确定它的首字根和尾字根，依次按下对应的字母键，首尾字根相同的十个汉字就显示在屏幕的下方，用“>”键和“<”键向后或向前寻求。找到时，按相应的数字键，选中的汉字就输入计算机了。所谓快速输入法，实际上就是首尾码的快速输入法。

首尾码输入汉字的主要优点是：见字识码，会写就会输入。

主要缺点是：重码很多，输入速度慢。

1.4.7 五笔字型

《五笔字型》汉字输入法是目前国内影响较大、应用面很广的一种高速汉字输入法。该方法又分为三种类型：简易型、普用型、高效型。简易型名为五笔画输入法，它采用横、竖、撇、捺、折五种基本笔画来概括组成汉字的各种笔画。输入时，只要输入汉字的开始四笔和最后一笔即可。普用型名为五笔桥输入法，它采用几十个字根和单笔画来概括组成汉字的各种笔画。它的基本输入方法是输入汉字的首笔对应的字根和其他部分的四个单笔画。高效型名为五笔字型输入法，它使用 130 个字根来概括组成汉字的各种部件。基本输入方法是输入汉字的前三个字根和最后一个字根。三种方法均有词组输入功能。

五笔字型输入法的主要优点是：见字识码，会写就会输入；输入速度很快，经过短期训练的人很容易达到 100 字/分钟以上的速

度。

主要缺点是：不会写的字和汉字库中的非汉字字符不能输入。

1.4.8 《大众码》汉字输入法

《大众码》汉字输入法也是目前国内应用十分广泛的一种快速汉字输入方法。它采用 128 个基本字元和 44 个联想字元来概括组成汉字的各种部件，共选定了 174 个字元。基本输入方法是输入汉字的前三个字元和最后一个字元。输入一个汉字的平均击键次数（包括空格键）为 3.42 键/字。

《大众码》输入法的主要优、缺点和五笔字型基本相同。

1.4.9 《自然码》汉字输入法

《自然码》汉字输入法是一种高度灵活的开放式汉字编码输入法，目前在国内的应用日趋广泛。自然码全码共五位，分为表音（声部及韵部）、表意（义部）、和表形（形部和尾部）三个独立的部分。在使用中可选用其中不同的 2 至 4 位搭配组合成字或词的输入码。由此演变出很多种输入方法。其特征为：音形结合，音里有形，形里有音，以音代形，以形表音，形音共存。

《自然码》的优点：简单易学、易掌握；输入速度快且汉字和非汉字图形符号均可输入；使用时，疲劳程度低。

没有明显的缺点。

1.4.10 综述

显然，任何一个使用计算机的人都没有必要掌握所有的汉字输入方法。一般来说，只要掌握其中一两种方法就足够了。但所有在计算机上进行中文处理的人员（使用《自然码》汉字输入法的人员除外），都应该掌握区位码、国标码、电报码这三种汉字输入方法的其中一种。特别是区位码汉字输入法。以便进行汉字输入时，万

一遇见找不到的个别汉字或用到非汉字字符,可用查表的方法找到。

在计算机(主要是微型机)上进行汉字输入,关键是微型机必须配备一个汉字操作系统(包括一个汉字库)。汉字操作系统有各种型号和版本。汉字库是把通常用的汉字聚集在一起,以备进行汉字的输入输出处理。目前,在国内使用的计算机汉字库一般都是按国标 2313-80 中规定的 7445 个字符组成。汉字库在微型机内部建立的方法有两种:软汉卡和硬汉卡。软汉卡是把汉字库放在硬盘或软盘上,启动汉字操作系统时调入微机内部。如:IBM-PC/XT, IBM-PC/AT 等微型机。硬汉卡是在生产微型机时(硬汉卡也可以单独加装),将汉字库直接装在微型机内部,如:长城 0520、长城 286 等微型机。不论哪种方法都必须和汉字操作系统联合使用。其中软汉卡是随着操作系统的启动,自动调入微型机内部的。

关于在计算机(主要是微型计算机)上选择汉字输入方法的操作(如何安装各种操作系统及汉字输入方法,这里就不赘述了,请参阅随盘带的安装说明),可大致分为两种:

1. 用“ALT+功能键”选择,如:IBM=PC/XT 等机型。
2. 用键盘上的方式键选择,如:长城系列微型机等。

第二章 五笔字型汉字输入法

§ 2.1 五笔字型概述

2.1.1

为了使用五笔字型输入法,在计算机的键面上,刻上五笔字型字根。这样,这个键盘就成了一个“字根键盘”,如图 2.1 所示。

键面上有“金木水火土,日月山石田,人口手”等字根。用户可以用这些字根象小孩子搭积木一样,拼合成许多汉字。如机器已经置在五笔字型的状态,用“日”和“月”左右拼起来就成为“明”,用“木”与“土”就可以拼合成为“杜”和“桂”。你只要在键面上轻轻地“点”一下有关字根,有关字根就会自动组合成你所要的汉字,并显示在屏幕上。

其中的奥妙也很简单。当你按了一个“木”字键时,计算机并没有如实地在机器中“写”进了一个“木”,而是记下了字根“木”所在键的英文字母 S。按了一个“土”键时,机器记下了字母 F。当你连续按了“木”、“土”时,机器记下了 SF。这时, SF 就成了“杜”字的代号(外码)。机器再按照这个代号,从存有几千个汉字图形的“字库”中,把“杜”字的图形调出来,显示在荧屏上,从而就完成了“杜”字的输入。

五笔字型方案采用“字词兼容”技术。在输入操作中,字与词之间没有什么换挡操作。当你连续按下“禾、石、讠、戈”四个字根时,屏幕上就会显示“科研计划”这个词。实际上,机器记下了与“禾、

金 金 金 35Q	人 人 人 34W	月 月 月 33E	向 向 向 32E	采 采 采 31T	第 第 第 41Y	江 江 江 42U	水 水 水 63I	火 火 火 44O	之 之 之 48P
工 工 工 15A	米 米 米 14S	大 大 大 13D	士 士 士 13P	五 五 五 11G	圖 圖 圖 31H	巨 巨 巨 22J	口 口 口 23K	田 田 田 24L	； ； ；
乙 乙 乙 35X	名 名 名 35X	又 又 又 54C	女 女 女 53V	子 子 子 52B	己 己 己 51N	山 山 山 25M	√ √ √	∧ ∧ ∧	！ ！ ！

图 2.1 字根键盘

石、讠、戈”这四个字根相对应的一组字母，即 TDYA，也就是这条词汇的外码。机器根据这组外码，在“词汇库”里把它对应的四个汉字的图形一次调出来。操作的过程和结果联系起来看，就好像是：取“科研计划”一词中每个汉字的第一个字根，就可以组成四个完整的字。其实质是人为地在两者之间建立了对应关系而已。

为了加强这种直观印象，下面再举几例。

例1：S键上有“木”，B键上有“子”，按SB（再加按一个空格键）就组出了“李”字。

例2：W键上有“人”，按WWW（再按一下空格键）就组出了“众”字。

例3：I键上有“讠”，P键上有“宀”，X键上有“纟”，D键上有“石”，按IPXD就组出了“汉字编码”（即打每一字的第一部分）。

这个过程叫“拼形组字（词）”。据此，五笔字型可把成千上万的汉字和词语输入计算机，这叫“拼形输入”。

2.1.2 汉字的三个层次

汉字是一种象形文字。最古老的汉字只求图形相似而不计笔划曲直多少。秦始皇统一中国之后，车同轨，书同文，统一使用的篆书，还没有从图画文字中蜕变出来。只要画成一定的形状，多一“笔画”，少一“笔画”都没有关系。直到后来，汉字楷化之后，才对汉字图形的线条加以规范化，其笔势有了一定之规，形成了“笔划”。可见，笔划是后来确定的。

一个完整的汉字，并非笔划的线性排列，也非笔划的任意堆积。由若干笔划复合连接交叉所形成的相对不变的结构，叫做“字根”。将字根按一定的位置关系拼合起来，就构成了汉字。

可见，汉字划分为三个层次：笔划、字根、单字。

《五笔字型》方案要求：不考虑读音，不把汉字全部肢解为笔划，而是遵从人们的习惯书写顺序，以字根为基本单位来组字编