

兰斌明 编

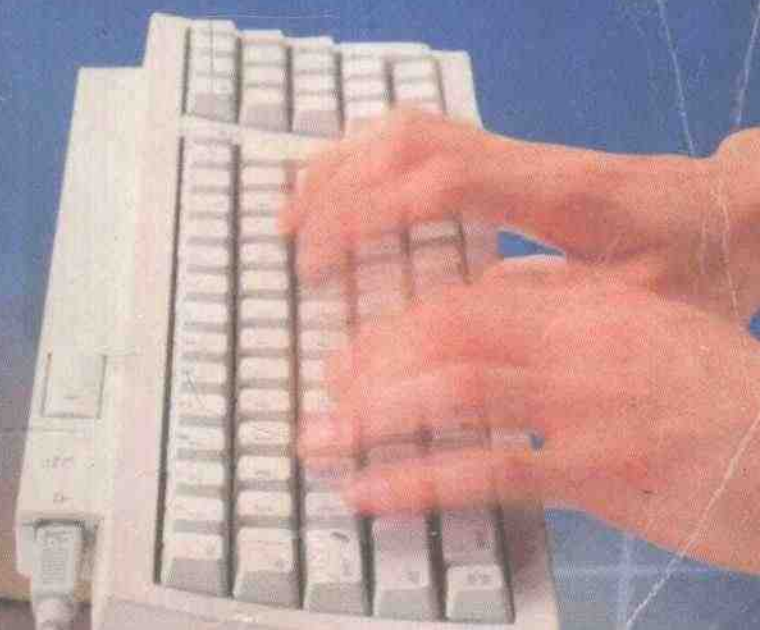
最新

ZUIXIN XIUZHEN WUBI
ZIXINGSUCHA SHOUC

袖珍五笔字型速查 手册



陕西科学技术出版社





袖珍五笔字型速查手册

兰斌明 编



陕西科学技术出版社

图书在版编目(CIP)数据

最新袖珍五笔字型速查手册/兰斌明编, —西安: 陕西科学技术出版社, 2002. 3

ISBN 7-5369-3459-9

I. 最… II. 兰… III. 汉字编码, 五笔字型—手册
IV. TP391.14—62

中国版本图书馆 CIP 数据核字(2002) 第 006445 号

出 版 者	陕西科学技术出版社 西安北大街 131 号 邮编 710003 电话(029)7211894 传真(029)7218236 http://www.snstip.com
发 行 者	陕西科学技术出版社 电话(029)7212206 7260001
印 刷	陕西宏业印务有限责任公司
规 格	890 mm × 1 240 mm 64 开本
印 张	4.125 印张
字 数	159 千字
印 数	1—5 000
版 次	2002 年 4 月第 1 版 2002 年 4 月第 1 次印刷
定 价	6.00 元

(如有印装质量问题,请与承印厂联系调换)

使用说明

五笔字型输入法是王永民先生于 1986 年发明的汉字输入法,1998 年王永民先生又推出了“98 规范王码”。98 王码编码科学、部件规范、好学易用。五笔字型输入是王码系列输入法中的一种,本手册所说的王码,仅指五笔字型。

为了使广大读者能够更快、更准确地掌握五笔字型输入法,使读者遇到不会输入的汉字时能够快速找到答案,为了照顾不同读者群的使用方便,本手册采用了拼音和部首两种查字法。

本手册还给出了金山双拼输入法编码。

本手册收集了 GB2312—80 的 6 763 个全部汉字,单字按汉语拼音的顺序排列,提供了五笔字型 86 版和 98 版单字编码。

本手册有以下特点:

可以从一个汉字的汉语拼音索引查出该汉字的页码,再查出五笔字型编码。

可以从一个汉字的部首目录中查出该汉字的部首,从检字表中查出该汉字的页码,难检字可以从难检字笔画索引中查出其页码,再查出五笔字型编码。

用小写字母排出二级简码和三级简码中可以省略不打的编码。简码是输入该字最少需要的字根码。

由于五笔字型 86 版和 98 版均有不同的版本,读者可

能会发现少量的汉字编码与本手册不同,但不同版本的编码规则是一致的,所拆分的字根是一样的。

由于本手册只给出少数常用的多音字,多数不常用的多音字在本手册中只出现一次,所以用部首查字法能更快地找到多音字。

我们衷心希望本手册成为五笔字型打字者和学习者的必备工具书,并期待读者提出批评和建议,以使得本手册更为合理和完善。

编者

2002 年 1 月

汉字编码与五笔字型输入法

1. 汉字编码

汉字编码有“内码”和“外码”两个概念。内码是指在计算机内部进行存贮传递和运算所使用的数字代码。外码是通过计算机的键盘按照某一方式进行汉字输入时,机器与人进行信息交换所采用的汉字字形的符号。汉字的输入方式可以不同,但是对于同一个汉字来说,它的内码却是相同的。五笔型、区位码、智能、全拼、双拼、电报码等录入编码都是外码。

计算机汉字处理的最关键和最基本的问题是解决汉字与计算机内机器码的对应关系。计算机中任何一个字符信息都是由一个二进制数表示的,这种关系是一一对应的、有序的。英文字符采用的是 ASCII 码 (American Standard Code for Information Interchange, 美国信息交换码), 就是西文的机内码, 它使用一个 8 位的二进制数, 也就是一个字节来表示英文字母、数字及一些特殊符号。8 位二进制数总共可以组成 $2^8 = 256$ 个不同的数值, 对应 256 个字符。

计算机中汉字代码一般采用两个字节表示, 共可以表示 $256 \times 256 = 65\,536$ 个汉字, 第一个表示区, 第二个表示位。这些汉字代码就是汉字内码。1980 年我国制定了 GB2312—80 字符集, 国标码的区号和位号都是 A1 ~ FE (16 进制)。这是我国第一个简体汉字内码的国家标准, 包括

6 763个汉字和 700 多个符号。国标码分成 94 区,每区 94 位, $94 \times 94 = 8\,836$ 个编码。1~15 区是字母、数字、符号;16~55 区是一级汉字字库,56~87 区是二级汉字字库,88~94 区是补字区。1990 年我国又制定了 GB12345—90,常称为 GBFT(国标繁体)。

此外,港台地区及海外华人一般都使用繁体汉字,BIG5 码是台湾计算机界实行的汉字编码,也称它为“大五码”。为方便繁简体汉字信息的转换与交流,BIG5 推出简体标准即 BIG5(JT)码。尤其近几年亚洲地区的交流日益密切,汉字内码的统一也日渐迫切。中国、日本、韩国又共同建立和推出了《CJK 统一汉字编码字符集》,即国标 GB13000.1,它完全等同于国际标准《通用多八位编码字符集(UCS)》,即 ISO/IEC10646.1—1993,UCS 是世界上各种语言的书面形式以及附加符号的表示、传输、交换、处理、存储、输入及显现的通用标准。其中从 0x4E00~0x9FA5 的连续区域内包含了 20 902 个来自中国(含台湾地区)、日本和韩国的汉字,称为 CJK(Chinese Japanese Korean)汉字。显然,CJK 是 GB2312—80、GB12345—90 及 BIG5 等字符集的超集。我国的蒙文、藏文、彝文作为未来扩充文种列入 UCS 的附录 N 中。

为解决 GB2312—80 的汉字收字不足、繁简汉字不能同平面共存、简化代码体系间不能转换等瓶颈问题,我国又制定了汉字扩展内码规范即 GBK,它是包含 21 003 个汉字的繁简体共存的大字符集国家标准。在保持已有软件兼容

的前提下,向最终的国际统一双字节字符集标准推进。GBK 制定的原则是与国家标准 GB2312—80 的内码体系标准完全兼容,并在字汇一级支持 ISO10646.1 及国家标准 GB13000.1 的全部 CJK 汉字,其非汉字符号同时涵盖大部分常用的 BIG5 非汉字符号。作为非 ISO10646.1 体系的代码页(Code Page),适用于中文信息的处理、交换、存储、显现、输入和输出。该规范的正式中文名称为“汉字内码规范”,英文名称为“Chinese Internal Code Specification(直译为:汉字国际内码规范)”,简称 GBK。GB 代表国标,K 是“扩展”一词汉语拼音的第一个字母,所以也称为“汉字扩展内码规范”,以区别于原来的内码规范 GB2312—80。在 Windows 95/98 中文版的“Windows”目录下,有一名为 GBK.TXT 的文件,它包含有 GBK 代码集的全体字符,可使用“书写器”打开并浏览其中的汉字与符号。

GBK 作为非 UCS(ISO10646.1)体系的代码页(Code Page),适用于中文信息的处理、交换、存储、显现、输入和输出,内容包括:

字汇 GBK 包含有 GB2312—80 的全部汉字及非汉字符号,包含有 GB13000.1 中的其他 CJK 汉字(以上 2 项合计为 20 902 个国际化汉字),包含有《简化字总表》中尚未收入 GB13000.1 的 52 个汉字;还包含我国《康熙字典》及《词海》中尚未收入 GB13000.1 的 28 个部首及重要构件;还收入了从 ISO10646.1 和 GB13000.1 的 CJK 兼容区挑选出来的 21 个汉字,计 21 003 个汉字、部首及构件。

字序 GBK 处理字序时的规则是 GB2312—80 中的汉字仍然依照原有的一级汉字和二级汉字,分别按拼音、部首与笔画排列;GB13000.1 中的其他 CJK 汉字,按照代码大小顺序进行排列;追加的 80 个汉字、部首和构件,与上述 2 类字汇分开,按照康熙字典的页码与字符单独进行排列。

码位分配 GBK 在码位分配时(图 1),总体上采用了 8140 ~ FEFE 的矩形区域,并剔除了 0x7F 的一条线,共有 23 940 个码位。汉字区共有 21 008 个码位,GB2312—80 的汉字分配在 B0A1 ~ F7FE 区,有 6 768 个码位,包括 6 763 个汉字。图形符号区共有 1 038 个码位。用户自定义区共有 1 894 个码位。

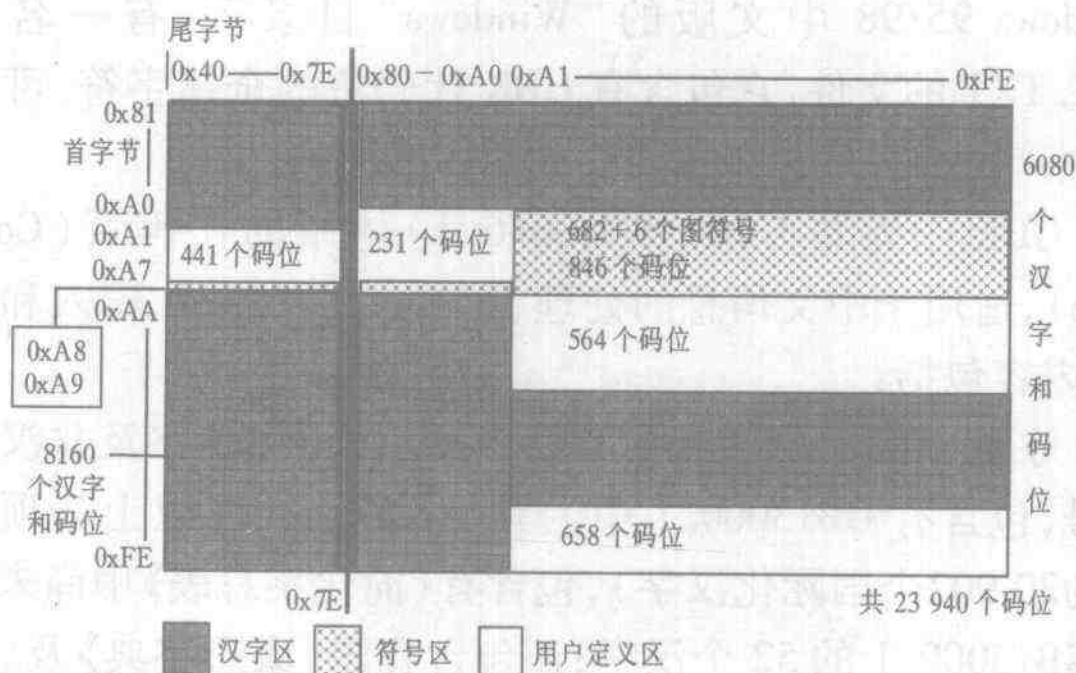


图 1 GBK 的码位分配表

2. 五笔字型输入法

五笔字型输入法是王永民先生于1986年发明的汉字输入法,它包括了GB2312—80的6 763个汉字,以字型进行编码,是目前计算机中最为成熟和最受欢迎的汉字输入法。五笔字型以86版王码五笔字型输入法为开端,经过十余年大量社会实践的检验,已成为国内占主导地位的汉字输入技术。1998年王永民先生又推出了“98规范王码”。98版王码编码科学、部件规范、规则简明、好学易用。目前各种汉字操作系统和Windows环境下都装有五笔字型输入法。

98 王码的优点

在86版五笔型输入法中,编码的基本单位称为字根。字根是指若干笔画复合连接交叉所形成的相对不变的结构,字根是构成字的基本单位,86版五笔字型的字根有125个。

在98规范王码的部件选取中,笔顺符合国家语言文字规范,86版中需要“拆分”的许多笔画结构,如“夫、甫、气、羊、母、丘、皮、毛、戊、艮”等,在98王码中都不需要拆分了,可以整体编码,再也不会为许多字的拆分而感到困惑了,所以易学易用。

在98王码的系列软件中,除了98版五笔字型输入法外,还包括王码智能拼音、规范五笔画、王码音型输入法等多种输入法。这些输入法是为了满足哪些对输入速度要求不高,希望很快掌握它的用户而设计的。

98 王码能处理多种字集。内码是在计算机内部表示汉字和各种符号的机内编码,不同的中文系统往往采用不同的内码标准,中国使用国际码(GB 码),台湾地区使用大五码(BIG5),其他国家和地区还采用 TCA 码、CNS 码、IS 码和 IBN5550 等机内编码标准。98 规范王码系列软件能够处理多种字集。

98 王码系列软件的新增功能

自动造词功能是王码智能拼音输入法所提供的一种造词方式,除此以外,用户还可以在编辑文本的过程中,从屏幕中取字造词。用屏幕取字造词方法造出来的词,还可以供其他输入法共享使用。

为了提高录入速度,王码智能拼音输入法还专门提供了重码动态调序功能。重码动态调序功能即能根据用户的输入法内容,自动计算字词的使用频率,把使用频率高的字或词调整到重码区的前端,以减少翻页。

码表编辑器是 98 王码系统为高级用户提供的另一个强有力的实用工具。利用该功能,用户可以对 Windows 95/98 中 GBK 字符集的 21 003 个汉字的五笔型编码和五笔画编码直接进行编辑修改,还可以建立自己的容错码。

98 王码软件克服了许多中文平台软件产品互不兼容、自立标准等缺点,立足于多内码实时转换、动态翻译的新技术。

五笔字型输入规则

1. 键名汉字输入

把键名汉字所在的键连击四下,如:

王 GGGG 白 RRRR

键名汉字所在键如下表所示:

金 Q	人 W	月 E	白 R	禾 T	言 Y	立 U	水 I	火 O	之 P
工 A	木 S	大 D	土 F	王 G	目 H	日 J	口 K	田 L	
学习键 Z	么(乡) X	又 C	女 V	子 B	己 N	山 M			

2. 成字码元输入

在 98 王码中,将汉字编码的单位定为码元(在 86 王码中称为字根),除键名以外自身也是汉字的码元(字根)输入方法为:

- (1)报户口即击一下成字码元(成字字根)所在键。
- (2)击第一笔所在键。
- (3)击第二笔所在键。
- (4)击最后一笔所在键。

注:不足 4 画者补击空格键。

	报户口	第一笔	第二笔	最后一笔
石:	石	一	丿	一
厂	厂	一	丿	空格键

3. 合体字的输入

(1) 合体字输入的取码规则

书写顺序 一般情况下按照正确的书写笔顺进行,如:
“新”字取码为:立、木、斤。

取大优先 按书写顺序取码时,尽可能取笔画多的码元(字根),如“估”取码为:亻、古,不能拆分为亻、十、口。

兼顾直观 在取码时,为了使码元特征明显、易辨,照顾字根的完整性,放弃前两个规则,如:“国”字取码为:口、王、丶。

能连不交 当一个字既可以视作“相连”的几个码元(字根),也可视作“相交”的几个码元(字根)时,取“相连”的码元(字根),这样更能显示笔画结构的特征。如:

天:一、大(两者是相连的) (正)

二、人(两者是相交的) (误)

能散不连 汉字的笔画和字根、字根与字根之间的关系,可以分为“散”、“连”、“交”三种关系,因此码元之间也有这样三种关系,如:

倡:“亻”、“日”、“日”之间是散的关系。

自:“丿”、“目”之间是连的关系。

夷:“一”、“弓”、“人”之间是交的关系。

当一个汉字的几个码元之间的关系,既能按“散”,又能按“连”取码时,规定只要不是单笔画,一律按“能散不连”判别,按“散”的关系取码。如:

午:丿 十(按“散”取码) (正)

丿、干(按“连”取码) (误)

(2) 合体字的输入

按码元(字根)数量的不同,合体字又可分为多元字、四元字、三元字和二元字四种。

多元字的输入 多元字指有四个以上码元(字根)的字,其输入时按照笔画的书写顺序,将第一、第二、第三及最后一个码元(字根),即“一、二、三、末”共编四个码元(取四个字根),如:

额:宀、夕、口、贝(PTKM)

四元字的输入 四元字指恰好有四个码元(字根)特征的字,其输入时按照笔画书写顺序将四个码元(字根)依次编码(取字根),如:

蝉:虫、乚、日、十(JLJF)

二元字和三元字的输入 二元字指有两个码元(字根)的字,三元字指有三个码元(字根)的字,凡是取不够四个码元(字根)的汉字输入时,需加一个“末笔字型识别码”,如还是不足四个码元(字根)时,则补打空格键,如“洱”字的码元(字根)为“氵”、“耳”,其所在键位为“I”、“B”,还需加一个识别码“G”,最后补打空格键。

末笔字型识别码 末笔字型识别码是为了区别码元(字根)相同,字型不同的汉字而设置的,只适用于不足四个码元(字根)组成的汉字。汉字共分三种字型:左右型、上下型、杂合型。在五笔字型中,笔画分为五种:一、丨、丿、㇏、乙。字型与末笔交叉的可能性有15种,其识别编码见

识别码表。

识别码表

字型 末笔型	左右型	上下型	杂合型
横	G	F	D
竖	H	J	K
撇	T	R	E
捺	Y	U	I
折	N	B	V

4. 简码输入规则

为了提高汉字的输入速度,对常用汉字只取最前的一至三个码元(字根),再加空格键输入即可,称简码输入。简码共分三级。

一级简码 “一、地、在、要、工、上、是、中、国、同、和、的、人、主、产、民、了、经”的一级简码击该字的第一个码元(字根)所在键再加空格即可,其余一级简码汉字人为规定了它的键位,一级简码共 25 个:

一(G) 地(F) 在(D) 要(S) 工(A)
 上(H) 是(J) 中(K) 国(L) 同(M)
 和(T) 的(R) 有(E) 人(W) 我(Q)
 主(Y) 产(U) 不(I) 为(O) 这(P)
 民(N) 了(B) 发(V) 以(C) 经(X)

二级简码 二级简码汉字的输入只要输入该汉字第

一、第二个码元(字根)的代码和空格即可,二级简码共有625个字。

三级简码 三级简码汉字的输入只要输入该字第一、第二、第三个码元(字根)和空格即可。

5. 词组的输入

双字词组 分别取两个字的第一、第二个码元(字根)的代码输入。如:

我们: 扌、扌、亻、门(TRWU)

合理: 人、一、王、日(WGGJ)

三字词组 取前两个字的第一码元(字根)代码,取最后一个字的前两个码元(字根)代码输入。如:

操作员: 扌、亻、口、贝(RWKM)

现代化: 王、亻、亻、匕(GWWX)

四字词组 取各个字的第一个码元(字根)代码输入。如:

调查研究: 讠、木、石、宀(YSDP)

中华民族: 口、亻、巳、方(KWNT)

四字以上的词组 取第一、第二、第三及最后一个字的第一个码元(字根)代码输入。如:

中国人民解放军: 口、口、人、宀(KLWP)

五笔字型 86 版与 98 版的比较

1. 键名字的不同

除 X 键的键名不同外(98 王码为“幺”,86 王码为“纟”),其余相同。

2. 成字码元(成字根)的不同

在 98 王码中将汉字编码的单位定为码元,除键名以外自身也是汉字的码元称为成字码元,而 86 王码则称为成字根,98 王码中增加的成字码元为:毛、夫、母、丘、羊、业、甫、戊、甘、皮、气、未、艮,但在 86 王码中的成字根“马”,在 98 王码中不是成字码元。成字码元(成字根):儿、力、广、乃、臼,在 86 王码和 98 王码中的键位不同。

86 王码成字根和 98 王码成字码元及所在键位见下表:

成字码元 (成字根)	86 王码	98 王码	成字码元 (成字根)	86 王码	98 王码
夕	Q	Q	儿	Q	Q
八	W	W	几	M	W
用	E	E	力	L	E
毛		E	臼	V	E
斤	R	R	丘		R
气		R	手	R	R
竹	T		方	Y	Y