

现代汉字 输入法

吴良占 沈美莉 编

浙江科学技术出版社

现代汉字输入法

吴良占 沈美莉 编

浙江科学技术出版社

内 容 提 要

本书介绍了目前几种最流行的用键盘输入汉字的方法，并附有编码表，可满足不同读者的要求。全书包括基本知识、输入方法、输入法的调用与安装、造字等几方面内容，其中输入方法从音码、形码和音形结合码等不同的编码形式进行讲述，还介绍了98标准王码这一最新的汉字输入方法，读者可根据自己的特点，选学其中一二。

书 名	现代汉字输入法
编 者	吴焱占 沈美莉
出 版	浙江科学技术出版社
印 刷	杭州富春印务有限公司
发 行	浙江省新华书店
制 作	浙江科学技术出版社计算机图书工作室
读者热线	0571—5157523
电子信箱	hzzjkj@public1.hz.zj.cn
开 本	787×1092 1/16
印 张	13
字 数	320 000
版 次	1999年1月第一版
印 次	1999年1月第一次印刷
书 号	ISBN 7-5341-1204-4/TP·67
定 价	21.00元
责任编辑	熊盛新
封面设计	金 晖

前 言

汉字输入是中文信息处理和办公自动化的前提，它的重要性是不言而喻的。

近年来，汉字输入技术发展很快，已从单一的键盘输入发展到多元化的语音输入、手写输入和扫描输入等。去年，中国科学院的汉王公司又推出了“汉王”系列，像汉王笔、汉王OCR、汉王全能阅读器、汉王听写输入系统、汉王读写听输入系统等，该系列采用了先进的汉字输入技术，不但让计算机“认”汉字、会“听”汉字，而且会“说”汉语、会“写”汉字。因此，有些人会认为汉字输入已不再需要键盘了。其实，这是片面的，因为任何一个新事物，都有一个产生、成长、壮大和完善的过程。就目前来说，由于多方面的原因，用键盘输入汉字仍然是汉字输入的主要手段。当然，我们也相信，将来计算机的语音识别、手写体识别一定会解决得很好，但用键盘输入汉字还是不能缺少，只不过不像现在这样重要而已。因此，我们编写了《现代汉字输入法》一书。本书语言简洁、通俗易懂，主要介绍了目前最流行的几种汉字输入方法，像智能拼音输入法、双拼双音输入法、五笔字型（包括最近推出的98标准王码）输入法、自然码输入法以及Windows 95/98下主要汉字输入法，并附有智能拼音、双拼双音、五笔字型和自然码等输入法的编码表，供用户在学习时查阅参考。

本书内容共分4章。其中，第1章和第2章的智能拼音及双拼双音输入法及第4章由吴良占先生编写；第2章的五笔字型输入法和自然码输入法、第3章及汉字编码表由沈美莉女士编写。由于作者时间紧张、水平有限，书中不足之处难免，欢迎广大读者批评、指正。

编 者
1998年7月

目 录

第1章 计算机汉字输入概述	1
1.1 计算机汉字输入与处理	2
1.1.1 汉字的字形表示与字库	2
1.1.2 汉字的编码与输入	4
1.2 汉字操作系统	7
1.2.1 汉字操作系统发展过程简介	7
1.2.2 UC DOS和Windows两个操作系统简介	8
1.3 汉字输入法简介	10
1.3.1 编码方案的划分	11
1.3.2 有关汉字输入的提示行信息	11
1.4 指法练习	12
1.4.1 打字姿势	12
1.4.2 手指动作与盲打	12
1.4.3 键盘的基本键与手指分工	13
第2章 汉字输入法	15
2.1 拼音输入法	15
2.1.1 全拼输入法	15
2.1.2 智能全拼输入法	16
2.1.3 双拼双音输入法	24
2.1.4 Windows95中文输入法简介	29
2.1.5 特殊符号的输入	33
2.2 五笔字型输入法	34
2.2.1 笔画、字根和汉字的关系	34
2.2.2 五笔字型字根键盘	38
2.2.3 汉字的拆分	39
2.2.4 末笔字型识别码	43
2.2.5 汉字的取码和输入方法	45
2.2.6 其 他	49
2.3 自然码汉字输入法	51
2.3.1 自然码的双拼编码规则	51
2.3.2 自然码的形义码规则	52
2.3.3 单字输入	54
2.3.4 词组输入	57
2.3.5 自定义词组	61
2.3.6 用自然码输入其他字符	64

2.3.7	其他功能	66
2.3.8	自然码输入操作功能简介	67
2.4	新版五笔字型输入法	68
2.4.1	关于码元	69
2.4.2	新五笔字型输入方法	70
第3章	汉字输入法的安装与使用	73
3.1	UCDOS汉字输入法安装与使用	73
3.1.1	UCDOS汉字输入法简介	73
3.1.2	UCDOS汉字输入法的安装	75
3.2	在Windows3.2下安装和使用汉字输入法	75
3.2.1	汉字输入	75
3.2.2	常规汉字输入法的安装	76
3.2.3	Windows版五笔字型的安装	76
3.3	Windows95下汉字输入法的安装与使用	77
3.3.1	汉字输入的使用方法	77
3.3.2	汉字输入的切换方法	78
3.3.3	汉字输入法的热键设置	78
3.3.4	汉字输入法的删除与添加	79
3.3.5	其他汉字输入法的安装	80
3.3.6	WM9801五笔字型新版本输入法的安装	81
第4章	造字	85
4.1	UCDOS的造字程序	85
4.1.1	轮廓造字程序MKFNT的功能与有关名词	85
4.1.2	MKFNT的功能键	86
4.1.3	使用MKFNT的造字过程	87
4.2	Windows95的造字功能	89
4.2.1	造字程序的进入与造字环境	90
4.2.2	工具的使用及编辑操作	94
4.2.3	怎样造字	96
4.2.4	使用新造的字	101
附录	国标一、二级汉字多种编码表	102
第十六区	a~bao	103
第十七区	bao~bing	104
第十八区	bing~chang	105
第十九区	chang~chu	107
第二十区	chu~dai	108
第二十一区	dai~die	109
第二十二区	ding~er	111
第二十三区	er~fu	112
第二十四区	fu~geng	114

第二十五区	geng~ha	115
第二十六区	hai~hu	116
第二十七区	hu~ji	118
第二十八区	ji~jian	119
第二十九区	jian~jin	120
第三十区	jin~jun	122
第三十一区	jun~kui	123
第三十二区	kui~li	125
第三十三区	li~long	126
第三十四区	long~man	127
第三十五区	man~mei	129
第三十六区	mo~ning	130
第三十七区	ning~pi	131
第三十八区	pi~qia	133
第三十九区	qia~qū	134
第四十区	qū~san	135
第四十一区	san~sheng	137
第四十二区	sheng~shu	138
第四十三区	shu~ta	140
第四十四区	ta~ting	141
第四十五区	ting~wei	142
第四十六区	wei~xi	144
第四十七区	xiao~xuan	145
第四十八区	xuan~yao	146
第四十九区	yao~yin	148
第五十区	yin~yü	149
第五十一区	yu~zha	151
第五十二区	zha~zheng	152
第五十三区	zheng~zhu	153
第五十四区	zhu~zuo	155
第五十五区	(一、丨、丿、丶、乙、厂、匚、冂、日、イ)	156
第五十六区	(亻、人、勺、勹、彡)	157
第五十七区	(彳、冫、讠、冫、冫)	159
第五十八区	(冫、勹、力、又、凵、厶、土)	160
第五十九区	(土、土、艹)	162
第六十区	(艹)	163
第六十一区	(艹、井、大、尢、九、扌)	164
第六十二区	(扌、戈、口)	166
第六十三区	(口、口、巾)	167
第六十四区	(巾、山、彳、豸)	168

第六十五区	(豸、夕、夕、夕、广、卜)	170
第六十六区	(卜、冂、斗、彡)	171
第六十七区	(彡)	173
第六十八区	(彡、宀、辶、冂、尸、弓、中、女)	174
第六十九区	(女、小、子、马、彡)	175
第七十区	(彡、玄、彡、王)	177
第七十一区	(、韦、木)	178
第七十二区	(木、犬、夕、车)	179
第七十三区	(车、戈、戈、瓦、日、贝、见、牛、手)	181
第七十四区	(手、毛、气、父、片、月)	182
第七十五区	(月、欠、风、爻、文、方、火、灬、户、衤、心)	184
第七十六区	(心、申、水、石、业、目)	185
第七十七区	(日、田、皿、钅)	186
第七十八区	(钅、矢、禾)	188
第七十九区	(禾、白、瓜、鸟、疒)	189
第八十区	(疒、立、穴、衤、疋、皮、耒、耳、西、页)	190
第八十一区	(页、声、虫)	192
第八十二区	(虫、缶、舌、竹)	193
第八十三区	(竹、白、自、血、舟、衣、羊、米、艮、羽、糸、麦、走、赤、豆、西)	195
第八十四区	(西、足、豸、角)	196
第八十五区	(角、言、雨、齿、鬣、隹、金、鱼)	197
第八十六区	(鱼、革、骨、鬼、食、影、麻、鹿、黑、鼻)	199

第1章 计算机汉字输入概述

从某种意义上说, 计算机应用就是信息处理, 对于我国也不例外。不过, 在我国, 信息处理是以汉字为主的, 即我们通常说的“汉字信息处理”。什么是“汉字信息处理”呢? 所谓汉字信息处理是指计算机对汉字信息进行加工、处理的过程。它被广泛应用于各行各业, 深受广大用户的欢迎。

汉字信息处理系统主要由汉字信息输入、汉字信息加工和汉字信息输出三部分组成。它包括硬件系统和软件系统两大部分。硬件系统除主机外, 通常还需配置汉字输入/输出设备、汉字字模库、汉字显示终端和汉字打印机等; 软件系统除了通常的操作系统与编译系统外, 还应有各种汉字服务程序。例如, 汉字输入输出程序和汉字字模制作程序。因此, 我们必须了解和掌握汉字的输入、显示和打印等有关知识。下面是计算机汉字信息处理系统结构示意图(见图1.1)。

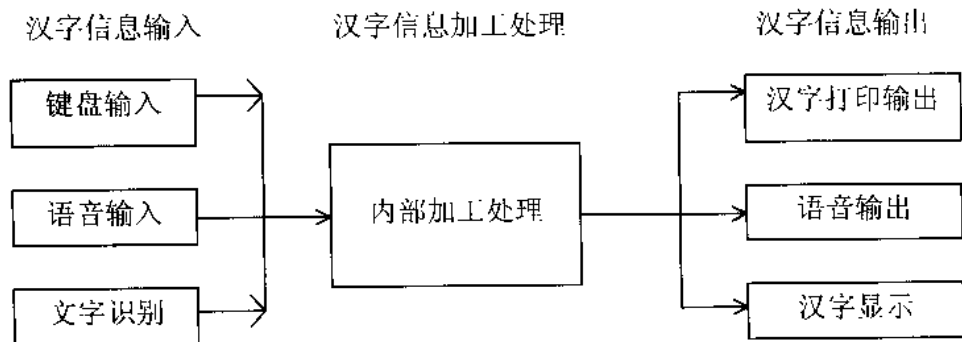


图1.1 计算机汉字信息处理系统结构示意图

由于全世界大约有1/4的人使用汉字, 而且我国的国际地位也在不断提高, 为了同世界接轨, 计算机汉化已势在必行。我们不但要让计算机“认”汉字和会“写”汉字, 而且还要会“听”汉语和会“说”汉语。为此, 中国科学院自动化所的汉王公司研究并推出了“汉王99”系列, 像汉王笔、汉王OCR、汉王全能阅读器、汉王听写输入系统、汉王读听写输入系统和汉王触摸屏等, 该系列旨在让汉语成为一种计算机通用语言, 因此, “汉王”系列不仅在国内是主流产品, 在我国港、台地区和东南亚也有很高的市场占有率, 其中:

汉王笔是一种手写输入中文、边写边输入的输入系统, 它支持中文、韩文和日文3种语言。其中的中文版可识别国标GB2312中的全部简体汉字6763个, 繁体一级汉字5401个, 我国香港地区异体字790个和大陆异体字、简化字13000多个。

汉王OCR是通过扫描来识别各种印刷品, 例如书本、报纸和杂志, 对文稿的识别率在99%以上, 识别速度也可达35~40字/秒。

汉王听、写输入系统则是把“汉王手写识别输入系统”与“IBM ViaVoice普通话连续汉语语音识别技术”集成在一起。“听”是指IBM语音识别输入, “写”是指汉王手写识别

输入,两者完美的结合,为广大用户提供了一个实用的、非键盘输入的汉字系统。

汉王读、听、写输入系统是汉王OCR3.1(印刷文稿、手写文稿)、汉王笔和汉王语音识别系统三种软件的集成。其中,“读”是印刷、手写文稿通过扫描识别,对扫描识别中的个别错误,可通过汉王笔进行修改、补充。

听写输入系统是一个相对独立的系统,不用扫描,也可以进行汉字录入。

在汉字信息处理系统中,首先遇到的是如何输入汉字问题。上面介绍的“汉王99”系列产品,好像汉字输入完全可以用汉王笔手写输入和用汉王听、写输入系统口述输入一样。其实,当前汉字输入还是要以键盘为主,因为“汉王99”推出还不久,要全面使用它还有一个过程。因此,本书仍介绍计算机键盘输入汉字的有关知识,为更好地使用汉字信息处理系统打好基础。

在使用计算机输入汉字时,大家自然会想到,输入几个字母或数字后怎么会出汉字呢?计算机又是怎样处理汉字输入呢?下面,就以上问题作一简单的解释。

1.1 计算机汉字输入与处理

1.1.1 汉字的字形表示与字库

许多西方国家,它们的文字都是用少量字母排列、组合而成。因此,它们只要通过一个比较简单叫做“字形发生器”的设备来产生这些字母及有关符号,再用这些字母和符号组成它们的文字。而我们的汉字却非常复杂,无论是拼音文字还是象形文字,都是以形来记录文字的音和意。因此,要在计算机中表示它们,就很难像西方国家那样用“字形发生器”来产生,必须设计一种特殊的数据结构,用它来表示汉字,权且称为“字形表示”,它把汉字以图形方式保存在计算机中。

1.1.1.1 汉字字形表示的实现方法

汉字字形表示方法,一般可以分为两类:点阵式字形和矢量式字形。点阵式字形不依赖于汉字本身的结构,还原技术简单,速度快,且便于和其他文字符号混用;缺点是占用大量的存储空间。矢量式字形是将一个汉字字形看成由许多直线段组成,这些直线段称为笔画(与汉字笔画不是同一含义),而这些直线段则用矢量来表示,这样,一个汉字就可用一组矢量信息来表示,用矢量表示汉字比用点阵表示汉字可使汉字表示信息大大减少。不过,在输出汉字时,仍须把矢量还原成点阵。下面简单介绍用点阵来表示汉字。

点阵法字形表示是:先将一个个汉字或字符在二维平面上用点阵表示出来,然后,将点阵数字化,即在有点处用“1”表示,无点处用“0”表示,再用一个二进制数表示每一行,这样,便得到了一个汉字或字符的二进制表示信息,称为字模。例如,对于一个 16×16 的点阵,就是把一个方块横向分成16格,纵向分成16格,共有256个小方格,即该矩阵有256个“点”。点阵中每一个点可以有“黑”、“白”两种颜色。用这样的点阵描出汉字的字形,称为“汉字点阵字形”。点阵数越大,所表示的字形就越精确。

若用二进制的“1”表示“黑”点,用二进制“0”表示白点,就很容易用二进制数来表示点阵,这就称为点阵的数字化。下面以“天”字为例,用 16×16 点阵来表示它(见图1.2和表1.1)。

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0							
1									.							
2									.							
3									.							
4									.							
5									.							
6	
7									.							
8								.	.							
9								.	.							
10							.			.						
11						.					.					
12				.								.				
13			.										.			
14		.												.		
15																

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	0	0	0	0	1	1	1	1	1	1	1	1	1	0	0	0
1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
6	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
7	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0
9	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0
10	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	0
11	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0
12	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0
13	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0
14	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

图1.2 “天”字形点阵表示

表1.1 “天”字的字形数据

地址	存储数据(二进制)	16进制	地址	存储数据(二进制)	16进制
00	00001111 11110000	0F F8	16	00000001 01000000	01 40
02	00000000 10000000	00 80	18	00000001 01000000	01 40
04	00000000 10000000	00 80	20	00000010 00100000	02 02
06	00000000 10000000	00 80	22	00000100 00010000	04 10
08	00000000 10000000	00 80	24	00001000 00001000	08 08
10	00000000 10000000	00 80	26	00010000 00000100	10 04
12	01111111 11111110	7F 7F	28	00100000 00000010	20 20
14	00000000 10000000	00 80	30	00000000 00000000	00 00

1.1.1.2 字 库

为了显示或打印汉字，我们要把每一个汉字都用点阵表示出来，并把点阵转换成字模，然后，按一定的顺序存入计算机中，这就形成了汉字字模库，简称字库。

当需要显示或打印汉字时，计算机找出字库中该汉字存放地址，然后取出该汉字的点阵信息并送入输出缓冲区，供显示或打印使用。

通用的汉字字库有：16×16点阵（简易型）、24×24点阵（普通型）和32×32点阵（提高型）。当然，还有高质量的精密字库，像用于精密汉字编辑排版系统的128×128点阵字库。一般地，显示字库是16×16点阵的，普通打印机的字库是24×24点阵的。

字库是建立在存储介质上面的，那么，一个普通的字库要用多大的空间来存放它呢？从表1.1可以看出，存储一个16×16点阵的汉字需要32个字节，8836个汉字就需282752字节，约276K字节。由此可见，字库存储容量是相当大的。因此，在存储字库时，不能把字库存放在太昂贵的存储器中，也肯定不能因降低成本而影响汉字读取速度。为了解决成本和速度的矛盾，我们通常采用两级存储方式，即将汉字库存放在硬盘中，使用时再将常用汉字调入内存。这种存放于硬盘里的字库称为软字库。有软字库就有硬字库，硬字库又称汉卡，它是将汉字字形信息固化在EPROM里面，然后插入计算机的扩展槽内，加电源后就可直接从汉卡中取出汉字，因而，汉字处理速度快、占内存少，但价格较贵。

1.1.2 汉字的编码与输入

把汉字存放在计算机中，既要占用大量内存，又不便于处理。考虑到汉字只在输出时才有意义，所以，实际上在计算机内部存储或处理汉字都采用一种与汉字一一对应的计算机内部逻辑码（简称机内码或内部码）。但是，要用户熟记内部码、直接用内部码输入汉字，是相当困难也不现实的。为了方便用户输入汉字，系统为用户提供了多种容易学习掌握的外部输入码（简称外部码或输入码）。输入码介于人、机界面之间，不同的编码方案有不同的输入码。下面就介绍汉字代码的有关知识。

1.1.2.1 汉字的代码

常用的汉字代码有：汉字交换码、汉字输入码、汉字机内码、汉字地址码和汉字字形码等。

(1) 汉字交换码。

在当今的信息社会里，计算机系统之间要经常交换信息，因此，必须遵守相同的规程和约定，指定交换码的标准信息。像ASCII码就是世界通用的信息交换码，汉字交换码则是在使用汉字信息交换时必须遵守的一种约定。

我国于1981年颁布了国家标准《信息交换汉字编码字符集》（基本集GB2312~80），简称国标码。目前，我们的计算机系统中所采用的汉字信息交换码就是国标码。国标码与ASCII码属同一制式，可以认为国标码是扩展的ASCII码。每个ASCII码用7位二进制数表示，共可表示 2^7 个ASCII码，其中，94个为字符代码。国标码也是以94个图形字符为基集，每个汉字交换码由两个字节构成，每个字节使用7位二进制数，编码范围0100001~1111110共94个。

国标码中收录了6763个汉字，按汉字的使用频度分为两级，其中一级汉字3755个，属常用汉字，按拼音顺序排列；二级汉字3008个，属非常用汉字，按部首排列。除此之外，国标码还收集了682个非汉字图形字符，包括一般符号、数字、拉丁字母、希腊字母和汉语拼音字母等。国标码中的7445个汉字和图形符号按照其代码排列在94×94的码表中，这张码表纵向分成94个区，由代码的第一个字节标识，横向分为94个位，由代码的第二个字节标识。国标码的分布情况如图1.3所示。

区	位	第二字节	
		00 01...20 21...	7E 7F
第 一 字 节	00		1 2 3 ... 94
	20		
	21	1	非汉字图形字符 (682个)
	29	9	
	2A	10	空白区
	2F	15	
	30	16	一级字库 (3755个)
	57	55	
	58	56	二级字库 (3008个)
	77	87	
	78	88	空白区
	7E	94	
	7F		

图1.3 国标码分布图

最近，我国制定了GB15000~90《信息交换用汉字编码通用字符集》，它包括了我国大陆地区、港澳台地区等的简繁体汉字，及日本、朝鲜等国家的汉字文化圈内所用的汉字，共计23000个，每个汉字使用2个7位编码表示，需用不同的控制字符去调用。

(2) 汉字机内码(内部码)。

按理说，国标码可以作为计算机系统的内部码，正如ASCII码作为计算机系统的内部码

一样。但是，事实并非如此。其主要原因是由于我国计算机技术走的是中西文兼容的道路，我们的计算机系统在汉字方式工作时，充分利用了原西文系统的大量软件资源。

国标码认为是由ASCII码扩展而来的，在中西文兼容系统中严格区分ASCII码和国标码，在计算机内表示汉字时不能直接使用国标码。通常是将双字节表示的汉字标准交换码的第一、第二字节的高位置为“1”，作为机内码，取值与汉字标准交换码一一对应。汉字机内码与国标码的关系可用如下公式表示：

$$\text{汉字机内码} = \text{国标区位码}H + 8080H + 2020H$$

其中，加8080H，即为双字节的最高位置为“1”；加2020H，是为了与ASCII码一致，因为ASCII码中前32个为控制字符与非打印字符。

(3) 汉字输入码。

汉字输入码是为了方便人们将汉字输入计算机而编制的代码，因为要用户熟记整个内部码，所以直接用内部码输入汉字是相当困难也是不现实的。目前，汉字主要还是通过键盘输入，因此，要求汉字输入码的规则简单易记、操作方便、码位要短、输入速度要快和重码率低等特点。据统计，国内各种各样的汉字输入方法已有一千多种，但应用较为广泛的只有二三十种，例如，各种拼音输入法、五笔字型输入法、自然码输入法、表形码和国标区位码输入法。

(4) 汉字地址码。

汉字地址码是指汉字字库中存储汉字的逻辑地址码。无论是硬字库还是软字库，汉字信息都是按一定的顺序存放在存储介质上，是连续有序的，且与汉字机内码的排列次序一致。这样，通过汉字机内码就可以方便地查到汉字地址码，通过汉字地址码就可以方便的读取汉字。

(5) 汉字字形码。

在上面介绍汉字字形表示时，用点阵来表示汉字的字形，再把点阵表示数字化，这就是以点阵方式形成的汉字字形码。一个16×16点阵汉字字形码需要用32个字节来存储；一个24×24点阵的汉字字形码需要用72个字节来存储；一个48×48点阵的汉字字形码需要用288个字节来存储。点阵越大，信息量越大，要求字节数越多。

以上5种代码都与某个汉字或某个图形字符相联系，都表示某一个汉字或图形的数据信息，故称之为“汉字数据码”或“图形字符码”。

在汉字信息处理系统中，除了汉字数据码外，还有一种虽不表示具体的汉字数据信息，却能影响汉字数据的传送控制、格式处理或解释执行等的代码，这就是汉字控制功能码。

汉字控制功能码必须按照国家标准或国际标准进行设计，它包括控制串定义符、引导符、格式控制符、排版用的各种各样的控制符等。

1.1.2.2 各种代码之间的关系

汉字的输入、加工处理及输出过程以及汉字交换码、汉字输入码、汉字内部码、汉字地址码和汉字字形码之间的关系见图1.4。下面具体说明几种代码之间的关系。

第一步是用户通过输入设备，例如，键盘，将汉字输入码输入计算机，汉字输入法有很多种，用户可以根据自己的需要选择不同的输入方法。

第二步是计算机会根据不同的输入码自动寻找相应的输入字典或转换函数将其转换成相应的内部码。因此，无论我们采用何种输入码输入同一汉字，存入计算机中的都是与字形码严格一一对应的内部码。

第三步是计算机在整个汉字存储和处理过程中一直使用内部码。无论对哪种字符进行排列、组合、转移、翻译等的处理都使用内部码。

第四步是当要进行汉字输出时，再通过输出字典或转换函数将相应的内部码转换成对应的字形码的地址码，然后，根据地址码取出字形码供汉字输出用。

第五步是把上面取出汉字字形码经过编排后，送到相应的输出设备输出，例如，送显示器显示或送打印机打印。

汉字控制功能码和汉字数据码组成汉字数据信息流，它贯穿汉字输入、机内处理和汉字输出等整个汉字信息处理过程。

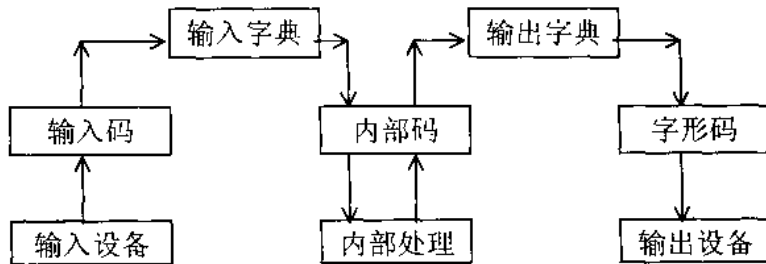


图1.4 输入码、内部码和字形码的关系

1.1.2.3 汉字服务程序

要使西文计算机具有汉字处理能力，实行中西文兼容，就必须附加汉字处理程序。主要内容有：

(1) 汉字输入程序：该程序是把汉字输入码输入到计算机中，然后，用汉字代码转换器和输入字典程序，把汉字输入码转换成汉字机内码。

(2) 读取汉字信息程序：该程序是把汉字机内码转换成相应的汉字地址码，再从字库（汉字字模库）中找出相应的汉字字形信息。

(3) 打印处理程序：是把从字库中读取的汉字字形信息从显示缓冲区送到打印机缓冲区的传送程序以及打印格式、字模尺寸变换等控制程序的集合。

(4) 汉字字模库的添加、修改、删除等程序。

(5) 汉字编辑程序、表格编辑程序等。

1.2 汉字操作系统

1.2.1 汉字操作系统发展过程简介

1983年，我国电子工业部第六研究所推出了第一个全部用软件实现汉字输入、打印的操作系统CCDOS 1.0和CCDOS 1.1版本，它可以在IBM PC机及其兼容机上运行。

由于当时的IBM PC机只有256KB的存储器，只有用区位码和国标码可以输入GB2312~80中的一、二级字库中的汉字，其他四种输入方法（拼音、音韵、电报、首尾）只能输入一级字库中的汉字，当时没有词组输入功能，而且，还有许多不足之处。

1984年,电子工业部第六研究所又开发出CCDOS 2.0和CCDOS 2.1两个版本,这两个版本有4种汉字输入方法,它们是拼音、区位、首尾和快速输入法。它们都能输入字库中一、二级汉字,并克服了许多早期的不足之处。它们不仅可以打印16×16点阵汉字,而且可以打印24×24点阵汉字,因此,受到了众多用户的欢迎,是早期最有影响的汉字操作系统。

1985年,他们又推出了以支持硬字库(汉卡)为特征CCDOS 3.0版本,它具有系统重组能力,大大增强了汉字输入输出功能,为用户腾出了更多的内存空间,使汉字的处理速度大大加快。

1986年后,汉字操作系统的发展进入高潮,先后出现了多种各具特色的汉字操作系统,例如,2.13A、UCDOS、LXCCDOS、CCDOS 4.0、SPDOS、WMDOS等。尤其是CCDOS 4.0,它代表了我国汉字操作系统发展的一个里程碑。1987年,电子工业部第六研究所采取了一系列新技术,汇集了当时国内汉字信息处理最新技术,特别是原来CCDOS的优点,把汉字输入方法扩充到了十几种之多,并提供了丰富的辅助输入手段。

由于受到设备的限制,且为了与西文软件兼容及编程简单,早期的汉字操作系统都是在西文操作系统的基本输入/输出模块BIOS基础上加上一层对汉字进行处理的外壳,把BIOS扩展成为CCBIOS,使西文和汉字均可以通过BIOS进行处理,完成汉字的输入、显示和打印。例如CCDOS 4.0、2.13(A~H)、UCDOS(1.0~2.2)、SPDOS(1.0~5.0)、WMDOS(1.0~5.0),它们都是通过扩展键盘管理中断模块、显示管理中断模块和打印管理中断模块的功能,实现汉字输入、显示和打印工作。

随着计算机软件和硬件技术的不断发展,对汉字处理的速度要求也越来越高,因此,必须采用一些发展起来的硬件新技术,不仅提供全面的中文BIOS,而且支持直接写屏,使西文软件能直接处理汉字,给用户带来很大的方便。同时,由于将显示字库存放于扩展内存里,缩小了整个汉字系统占用的内存空间,增加了用户的可用内存空间,这就要求系统整体优化设计技术相当先进。如UCDOS(3.0~5.0)、Windows 3.1中文版、Windows 95中文版都是全新的中西文兼容的操作系统,它们很快就受到广大用户的热烈欢迎,成为微机操作系统的发展方向。

1.2.2 UCDOS和Windows两个操作系统简介

优秀的中文操作系统很多,例如UCDOS、SPDOS、中国龙、天汇、联想、超想、工码、中文之星2.13等,这里不可能一一介绍,下面仅简单介绍UCDOS和Windows两种平台。

1.2.2.1 UCDOS操作系统

UCDOS是北京希望电脑公司的产品,该产品自1985年推出1.0版以来,经过不断加工、改进,先后推出了2.0、3.0、5.0、6.0、7.0等各种版本,功能日趋完善,拥有大量用户,是目前汉字操作系统中最具代表性的一个。

(1)系统的主要特色。

- 与设备无关的内核设计。系统的内核设计与外部设备完全独立,可以支持各种类型的显示卡、打印机、汉字输入方法和打印字库。还可根据不同类型的外部设备编制不同的驱动程序,例如,显示卡驱动程序、打印机驱动程序、汉字输入方法和打印字库接口程序等,用户可以根据需求选择相应的程序,充分发挥外部设备的作用。

- 支持直接写屏。西文软件毋需汉化就可进行中文处理,既保持了西文软件的面貌,又可显示和输入汉字。使用直接写屏和制表符识别技术,可以非常正确地识别出英文制表

符,使它与汉字共存于同一屏幕,而不发生冲突。

- 完备的汉字输入系统。系统提供了智能拼音、全拼、简拼、双拼、五笔、普通、自然码等十几种汉字输入方法。还提供了万能输入法管理器,用户可以根据需要方便地修改或编制汉字输入法。特别值得一提的是智能拼音输入法,它充分利用和发挥汉字拼音编码的规律,自动记忆词组、自动调整词频、模糊输入,大大提高了拼音输入效率。

- 强大的打印功能。系统提供的字库还原核心程序与字库格式无关,支持各种点阵字库、矢量字库和曲线字库,其他厂家的打印字库也只需编写一个简单的字库驱动程序即可在UCDOS下正常使用。系统还提供Windows True Type字库的驱动程序WinFnt,使Windows True Type字库可以在UCDOS下直接使用,它支持各种类型的打印机——针式、喷墨和激光打印机,还支持彩色打印机等。系统提供的独特的打印字库还原技术,其还原速度可与硬件媲美,使打印速度大大提高。由于点阵字库和曲线字库的有机结合,既可保证小字的打印质量,又可实现汉字的无级缩放打印,可在任意软件中打印 5120×8000 (点)的汉字。使用大字打印程序,最大可打印20平方米的汉字。

- 齐全的屏幕显示功能。系统提供了强大的图像显示和图像动态保护功能,屏幕图像不仅可以作为文件保存,也可以保存到XMS内存,图像文件的类型为标准的格式,图像显示可以无级缩放等。

- 造字功能。系统提供点阵字库和曲线字库的造字程序,可以利用现有的汉字很方便地造出新字。系统已经预造一百多个常用的非国标汉字。

- 双向英汉字典。系统提供了一个包含十多万条单字、词组、同义词、反义词的双向英汉字典。英汉字典常驻内存,可在任意环境下使用,自动显示英文单字的中文解释,也支持用鼠标在屏幕上取词。英汉字典也可以在汉化的Windows 3.X和Windows 95下运行,并率先创造了用鼠标右键取词方法,使字典使用更为方便。

其他的还有集成工作环境、计算机病毒防护,真正实现网络共享、名片管理等多项功能。

(2) 系统的启动与退出。

● 系统的启动。

系统提供两个启动的批命令文件,它们是UCDOS.BAT和UP.BAT。用UCDOS.BAT启动,不能使用汉字打印功能和WPS文字处理功能。用UP.BAT启动,由于加载了打印字库读取模块和汉字打印模块,因此,可以使用打印功能和WPS文字处理功能。

具体操作为:

开机后当出现DOS的提示符后,直接输入下面命令: C:\>UP或UCDOS,即可启动UCDOS。

● 系统的退出。

退出系统有两种办法。第一种办法为在DOS提示符下运行QUIT.COM文件,系统将退出UCDOS所占有的所有资源。第二种办法为直接按CTRL+F5复合键,即可退出UCDOS。

(3) 列出安装UCDOS后的目录结构。

下面以UCDOS 6.0为例进行说明。安装UCDOS 6.0后,所有的程序和数据文件都被拷贝到目标盘中,各类文件的目录分布如下:

C:\>(C盘根目录)