

PC-1500 微型计算机汉字系统 的初步研究

宛 公 展

(天津市气象科学研究所)

王 宝 成

(天津市塘沽气象台)

随着 PC-1500 微型机的普及, 大家对该机能否打印输出汉字提出了许多想法 (如编制汉字工资清单, 绘制各类中文表报等)。因此, 在 PC-1500 微型机上开发汉字, 研制中文系统软件, 已成为大家很感兴趣的一个课题。

本文提出一种完全靠软件 (用 BASIC 语言编写) 支持的汉字系统。该系统只占内存 1.3K, 它即可作为程序单独使用, 也可以当成一个子程序模块嵌入其它程序。当需要打印汉字时, 只需用规定的指令调用汉字子程序即可。汉字可以在任意方向 ($0 \sim 360^\circ$) 上打印, 其大小和字形 (长宽比例) 也可以随意控制 (参见图 1)。

一、用网格框架将汉字规格化

汉字是通过一笔一划书写形成的。其基本结构是由点、横、竖、撇、捺、钩、斜等基本笔划组成。但由于书写上的随意性, 以及基本笔划及其长短在字区中出现的随机性, 因此很难在内存有限的 PC-1500 微型机 (通常为

天津市气象科学
研究所计算机室
电脑中文系统

天津市气象科学研究所

天津市气象
科学研究所

天津市气象科学研究所

天津市气象科学研究所

天津市气象科学研究所

图 1 不同大小, 方向及不同高宽比例的汉字样品

8K)上实现由基本笔划拼接的汉字。但是,由于该机有很强的绘图功能,它的CE-150打印机实际上也是一部 $x-y$ 绘图仪,因此,完全可以将汉字看作是一种特殊的“图形”,用作图的方法打印汉字。

从图形的角度观察汉字,可将汉字笔划的结构简化为横、竖、撇(左斜)、捺(右斜)四种基本类型。通过控制这四种基本笔划的走向和长度,可以将绝大多数汉字“规格化”。规格化后的汉字具有类似仿宋体美术字的字形结构,字体美观,容易辨认。特别是,对规格化了的汉字,可以十分方便地进行编码,从而形成一种以笔划为基础的汉字信息代码。计算机读入这些代码后,便可以打印输出相应的汉字。

为了用尽可能少的编码指令打印出可以辨认的常用汉字,我们采用带有边框的 5×7 的网格,作为汉字四种基本笔划走向的框架。其中横向笔划(a)有 7×4 个,竖向(b)有 5×6 个,斜线(c)有 8×6 个,共106个笔划单元(见图2)。利用这个框架安排汉字笔划,可以完成大多数汉字的表示工作,只要笔划安排得当,汉字形体工整,容易辨认。图3是“中文汉字”四个字的规格化表示。

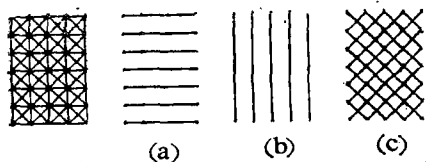


图2 汉字基本网格框架及其组成部分

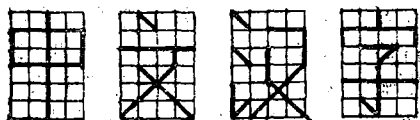


图3 “中文汉字”的规格化

二、规格化汉字的信息编码

将汉字安排在图2的基本网格上,即可进行信息编码。其原理是,框架上的每一个笔划单元,可看成是一个二进制代码,可用

“1”表示这个笔划“着墨”,用“0”表示“不着墨”,这样,即可显示出任意汉字图形。

为便于进行编码,先将框架的横线自下而上标记为1~7区,再将竖线从左至右标记为1~5区,斜线区从左至右标记为(1)~(4)

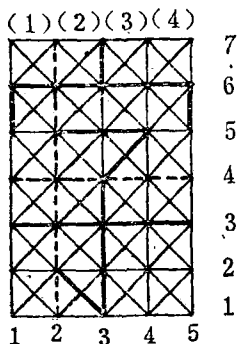


图4 汉字网格的分区

(见图4),并用十字线(图中的虚线)将框架上下左右各分成两部分。十字线相当于计算机穿孔纸带的中导孔,上与左为高字位,下与右为低字位。汉字笔划的信息编码指令规定为TDAB。其中, T 为笔形区分($T=1$ 为横线, $T=2$ 为竖线, $T=7$ 为撇, $T=8$ 为捺); D 为着墨区域区分(横线为1~7区,竖线1~5区,斜线为1~4区); A 为着墨编码(八进制)的高字位; B 为低字位。

比如,汉字的“字”,其编码情况可见图4,先编横划,第一横划出现在第6区,即横6区需全部着墨,二进制为1111,八进制编码为17(高字位为1,低字位为7);横5区仅中间需着墨,二进制编码为0110,八进制编码为06;类似地横3区编码为17。再看竖划,竖1区编码为20(二进制为010000,高字位为2,低字位为0),竖3区编码为47,竖5区编码为20;最后编斜划,2区右斜,编码为01(二进制000001,高字位为0,低字位为1),3区左斜,编码为10。

利用这种方法编码时,只需将画好的网格衬在较透明的白纸下面,然后在白纸上沿着网格的笔划单元书写汉字,写好后即可读出编码。其口诀是“一横、二竖、七撇、八捺”。上面的“字”,按TDAB指令编码格式写成完

附录：PC-1500 微型机汉字打印源程序清单(占 1.3 K 字节)

```

4:REM PC--1500"C
HINESE SYSTEM"
5:GRAPH :CLEAR
10:INPUT "M=?";M,
"DEG?";G,"H=?";
;H,"W=?";W:IF
G=0INPUT "P?";
P
15:DEGREE :G1=SIN
G:G2=COS G
20:INPUT "X0=?";X
0,"Y0=?";Y0:
GLCURSOR (X0,Y
0):SORGN
25:FOR I=1TO M
30:IF G=0THEN IF
(I-1)/P=INT ((
I-1)/P)THEN
GLCURSOR (0,-9
*H):GRAPH
40:FOR J=1TO 21
50:READ S:IF S=0
GOTO 70
60:GOSUB "C"
65:NEXT J
70:GLCURSOR (6*W*
G2,6*H*G1):
SORGN
80:NEXT I
90:END
100:"C":T=INT (S/1
000):D=INT (S/
100)-T*10
110:A=INT (S/10):A
=A-INT (A/10)*
10:B=S-INT (S/
10)*10
120:IF T>2LET F=D-
1:GOTO 220
130:IF A<>0LET E=0
:C=A:GOSUB "L1
"
140:IF B<>0GOTO 16
0
150:GOTO "E"
160:C=B
170:IF T=1LET E=3:
GOSUB "L1"
180:IF T=2LET E=-3
:GOSUB "L1"
190:GOTO "E"
220:IF A<>0LET E=3
:C=A:GOSUB "L2
"
230:IF B<>0LET E=0
:C=B:GOSUB "L2
"
240:"E":RETURN
300:"L1":GOTO 300+
C*10
310:U=2:V=3:GOTO 3
80
320:U=1:V=2:GOTO 3
80
330:U=1:V=3:GOTO 3
80
340:U=0:V=1:GOTO 3
80
350:U=0:V=1:GOSUB
"L":U=2:V=3:
GOTO 380
360:U=0:V=2:GOTO 3
80
370:U=0:V=3
380:GOSUB "L"
390:RETURN
400:"L2":U=2:X=3:
GOTO 400+C*10
410:U=1:Y=0:GOTO 4
80
420:U=2:Y=1:GOTO 4
80
430:U=2:Y=1:GOSUB
"L":U=1:Y=0:
GOTO 480
440:U=3:Y=2:GOTO 4
80
450:U=3:Y=2:GOSUB
"L":U=1:Y=0:
GOTO 480
460:U=3:Y=2:GOSUB
"L":U=2:Y=1:
GOTO 480
470:U=3:Y=2:GOSUB
"L":U=2:Y=1:
GOSUB "L":U=1:
Y=0
480:GOSUB "L"
490:RETURN
610:"L":IF T=1LET
Y1=D-1:Y2=Y1:X
1=U+E:X2=U+E:
GOTO "LE"
620:IF T=2LET Y1=6
-U+E:Y2=6-U+E:
X1=D+1:X2=X1:
GOTO "LE"
630:IF T=8LET X1=U
+F:Y1=U+E:X2=X
+F:Y2=Y+E:GOTO
"LE"
640:X1=X+T:Y1=U+E:
X2=U+F:Y2=Y+E
690:"LE":IF G<>0
GOSUB "RT"
700:LINE (W*X1,H*Y
1)-(W*X2,H*Y2)
710:RETURN
720:"RT":Z1=X1:Z2=
X2
730:X1=X1*G2-Y1*G1
:Y1=Z1*G1+Y1*G
2
740:X2=X2*G2-Y2*G1
:Y2=Z2*G1+Y2*G
2
750:RETURN

```

整的编码为：

1617, 1506, 1317, 2120, 2347, 2520, 7310, 8201, 0。最后一个 0 是这个字的结束符。计算机读到它,就会自动调整坐标原点,准备打印下一个字。

三、计算机程序软件的支持

对程序软件的要求是：根据读入的信息编码,进行解译,然后按用户的要求(如字体初始位置、大小、长宽比例、打印方向等),打印出相应的汉字。这个过程可用图 5 的流程图表示。程序设计的关键是找到对任何汉字都通用的划线(包括直线与斜线)基础坐标

系,其次是解决打印横与竖及撇与捺的高度兼容的译码系统,这样才能使程序设计大大精炼。至于汉字打印方向的控制,可利用坐标系旋转的原理,左乘一个正交矩阵来解决,即

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

这里, x, y 为旋转后的新坐标, x', y' 为原来的坐标。

PC-1500 微型机汉字打印源程序参见附录。

四、讨 论

本文介绍了仅用 1.3 K 内存,在 PC-

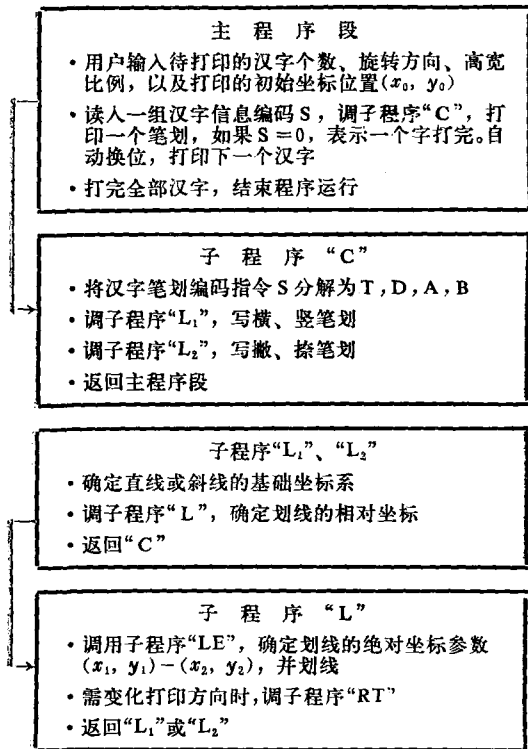


图 5 PC-1500 微型机汉字打印程序流程框图

1500 微型机上打印常用汉字的方案。打出的汉字字型工整, 易于辨认。汉字信息编码简单, 稍加练习即可学会。由于程序软件精炼, 打印汉字的速度也相对较快 (由于 CE-150 打印机本身性能的限制, 不可能非常快)。这对于只有 4—8K 内存的机器来说, 可以在不变更内存模块的条件下, 直接拿来使用, 而且不会影响该机的其它任何功能。特别是, 本文提出的汉字方案, 不对机器本身再规定什么“进入汉字状态”等, 因此, 打印汉字与其它各类数值计算, 都可以兼容在一个程序之中运行。另外, 该汉字系统的建立, 不需要增加任何硬件设备, 这对于一般用户是十分方便的。

如欲使打印的汉字更加美观与完善, 并能写出笔划复杂的不常见汉字, 可以变动图 2 的基本网格框架 (如改为 7×9), 并将程序中相应的译码系统稍加修改, 即可解决。有兴趣的读者不妨试一试。