

用分析综合法识别手写体汉字

顾志远 编译

本文介绍一种识别手写体汉字的分析综合法。该识别法的基础在于下述事实：汉字可以写成几种基本笔划的序列，而手写体汉字每一笔划又可按书写时笔点运动的相当简单的模型来描述。

该识别方法的作用过程如下：首先，体现笔点运动模型的形成过程，从输入文字图形中抽取笔划序列；然后，从笔划字典中查找所抽取的笔划序列来识别汉字。

该识别方法已经就点数为1000的手写体汉字进行了计算机模拟实验，正确识别率是82.4%。

* * *

引言 近年来，人们已经越来越多地注意汉字图形的自动处理。以前，主要致力于发展在线计算机输入和输出系统，各种方法均已取得一定的成效^(1,2)。但因为汉字数量庞大，结构复杂，至今，汉字自动识别对现有的信息处理技术来说，仍是一个困难的课题。此外，对于识别手写体汉字，还存在不同的书写者写同一汉字造成的各种变体的问题。

为了确定手写体汉字的识别方法，运用了下述事实：汉字可以写成几种基本笔划线段的序列，而手写体汉字每一笔划线段又可按照书写时笔点运动的相当简单的模型来描述。

若将这些基本笔划线段看作是识别的原始特征，则可提出一种使用分析—综合变化表的手写体汉字识别方法。这种方法的实行可分下述三步：

1. 预处理：对字体的大小和位置进行比例调整；将笔划宽度细化；及抽取初步特征。
2. 笔划抽取：从经过预处理的文字图形中用分析—综合法抽取基本笔划线段。
3. 识别：在包含有预定的手写体汉字的笔划字典中查找所抽取的笔划序列而确认汉字。

本文首先叙述为了获得手写的笔点运动的模型而进行的初步实验；然后说明该识别方法的步骤和功能；最后给出计算机模拟实验所得的手写体汉字识别结果。

笔点运动分析 为了研究书写时笔点运动的特性，设计了一种实验装置，并已投入运行试验⁽³⁾。该笔装置可作为计算机处理用的廉价图形输入装置。此笔装置产生三种电信号，分别比例于笔点在书写衬垫上移动的水平投影、垂直投影以及笔尖和衬垫间的压力。这些信号经过低通滤波、取样，并转换成数字形式；然后微分以产生笔点的速度和加速度信号。

欧洲文字书写时，笔点运动的节奏性很强。如图1所示。本文作者之一 (Eden)曾提出用某些三

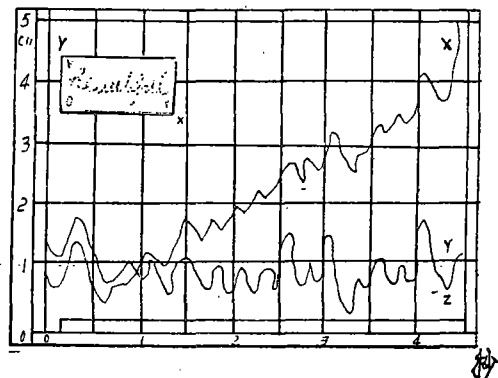


图 1

角函数作为欧洲文字书写时笔点运动的数学模型^{〔5〕}。Mermelstein 通过在笔划中采用加速和减速区域的概念,而扩展了模型^{〔6〕}。此外, Van der Gon 和 Mac Donald 也报告了对草体字的其它动态特征进行归纳的方法。

在如图 2 所示的手写汉字的情况下,笔点运动不如欧洲文字书写时那么有节奏。可以这样认为,因为汉字是固有的象形文字,在历史发展过程中,其字形是从美学观点考虑的,而不是从易于书写的角度考虑的。

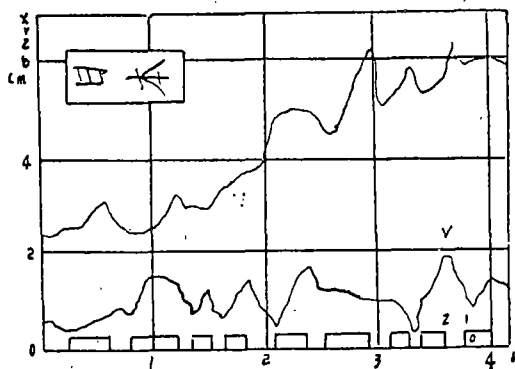


图 2

用我们的方法继续进行的分析发现,手写体汉字笔划中笔点的运动可以归纳为恒定的角加速度,在一个书写期间内,线速度按照抛物线变化,如图 3 所示^{〔9〕}。我们将实验结果加以模型化,提出描述汉字笔划中笔点运动的数学公式如下:笔点速度矢量 $\dot{V}(t)$ 由式(1)给出:

$$\dot{V}(t) = V(t) \cdot \exp i \theta(t) = [A(t-B)^2 + C] \cdot \exp i(\omega t + S) \quad (-t_1 \leq t \leq t_2, V(t) > 0, \omega = \omega_p(t > 0), \omega = \omega_m(t < 0)) \quad (1)$$

笔点的位移 $(x(t) \cdot y(t))$ 在 $t > 0$ 时由式(2), (3) 给出:

$$x(t) = \int_0^t V_x(\tau) d\tau + x_0$$

$$= \frac{1}{w_p} \left[A(t-B)^2 + C - \frac{2A}{w_p^2} \right] \sin(w_p t + S) + \frac{2A}{w_p^2} (t-B) \cos(w_p t + S) - \frac{1}{w_p} \left(AB^2 + C - \frac{2A}{w_p^2} \right) \sin S + \frac{2A}{w_p} \cos S + x_0 \quad (2)$$

$$Y(t) = \int_0^t V_y(\tau) d\tau + Y_0$$

$$= \frac{-1}{w_p} \left[A(t-B)^2 + C - \frac{2A}{w_p^2} \right] \cos(w_p t + s) + \frac{2A}{w_p^2} (t-B) \sin(w_p t + s) + \frac{1}{w_p} \left(AB^2 + C - \frac{2A}{w_p^2} \right) \cos S + \frac{2A}{w_p} \sin S + Y_0 \quad (3)$$

另外,对于 $t < 0$, 笔点位移同样可由与式(2)、(3)相似的公式给出。式中, (X_0, Y_0) 是 $t = 0$ 时笔点的位置; $V(t)$ 和 $\theta(t)$ 分别表示线速度的大小和方向; A, B, C, w, S 是参数。

已经通过实验检验了该模型的适用性,实验中,用上述关系式产生出手写体汉字的样品,并使样品确实符合关系式。其结果已经在早些时候发表。

用分析综合法的手写体汉字识别 Stevens 提出的分析综合法作为语言分析和识别方法是一

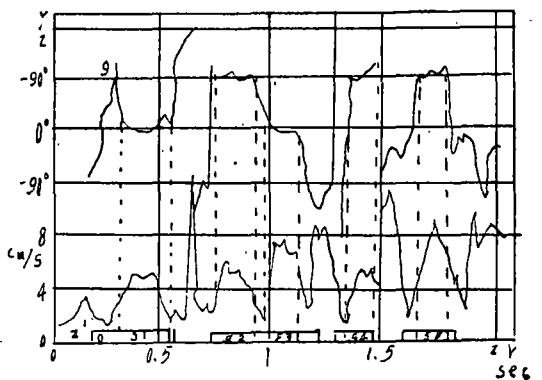


图 3

种人所共知的信息处理技术^[10, 11] (如图4所示)。该技术的特点是通过迭代反馈过程, 逐步逼近输入信号中包含的参数实际值; 而且, 该技术的目的还在于以研究有关输入信号产生的实际过程, 而改善分析的效能。

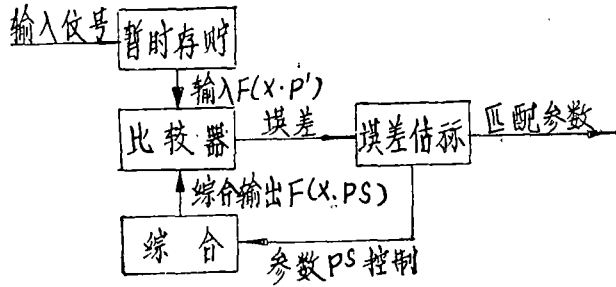


图 4

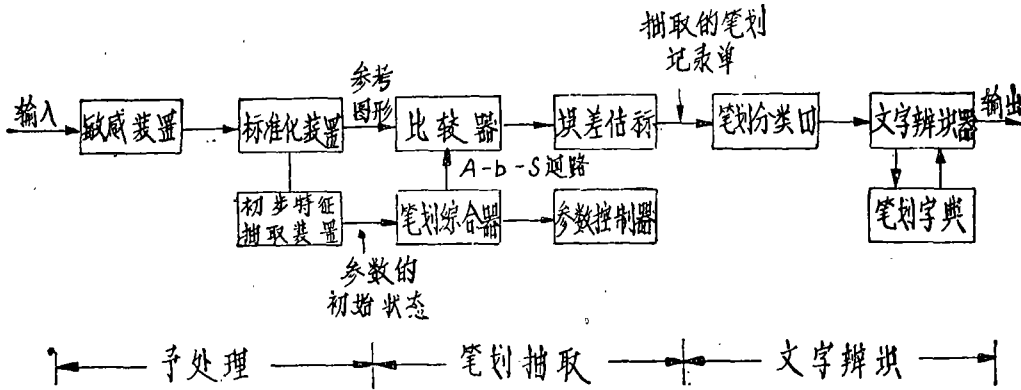


图 5

图5是该识别方法的方框图。该方法的特点在于利用分析综合法进行笔划抽取, 它又具体运用了上述分析所推导出的笔点运动模型。

预处理 用光导摄像机对输入文字图形扫描(在麻省理工学院所进行的后半部分工作, 曾用平板系统来搜集数据)。同时, 将扫描图片转换成 64×64 的图素阵列。每一图素取亮度值为1(黑)或0(白)。再就文字大小作比例调整及笔划宽度细化等, 进行标准化处理。然后, 将经过标准化处理的图形中的黑图素的位置以及它们与各自相邻的八个图素连结的方向存储在参考图形表中。预处理的最后阶段是初步笔划的抽取过程。在这一阶段, 已经标准化的图形中的黑图素用它们的互连关系来示踪, 其线性笔划线段(笔划主干)由最小方差近似法来提取。将提取出的笔划主干连同它们的长度、位置以及方向都存储在线性笔划线段表中, 再将这一些作为笔划综合的初始状态应用于下一次笔划提取的过程。图6中示出一个预处理的例子。

笔划提取 笔划提取过程包括笔划综合、比较、误差估算以及参数控制级。为了减少计算时间, 将上面所提出的笔点运动模型简化如下:

$$\langle \text{笔划} \rangle := \langle \text{点} \rangle | \langle \text{划} \rangle | \langle \text{点} \rangle \cdot \langle \text{拐} \rangle | \cdot \langle \text{拐} \rangle \cdot \langle \text{划} \rangle \cdot \langle \text{拐} \rangle \quad (4) \text{式}$$

(4)中, $\langle \text{点} \rangle$, $\langle \text{划} \rangle$, $\langle \text{拐} \rangle$ 的

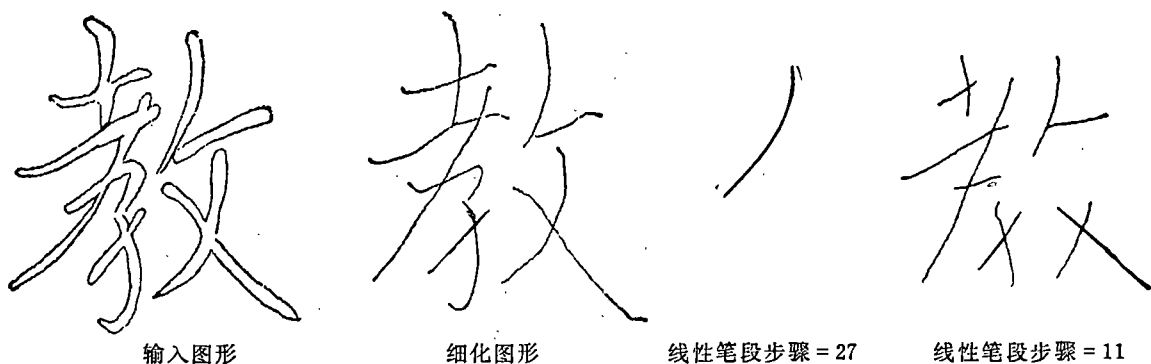


图 6

含义如下:

<点>: 图形中孤立的黑点;

<划>: 能满足式 (5)、(6)、(7) 的线性笔点运动。

$$\dot{V}(t) = V(t) \cdot \exp i \theta(t) = B \cdot \exp i s \quad (5)$$

$$\begin{cases} x(t) = Bt \cdot \cos S + x_0 \\ y(t) = Bt \cdot \sin S + y_0 \end{cases} \quad (6)$$

$$\begin{cases} x(t) = Bt \cdot \cos S + x_0 \\ y(t) = Bt \cdot \sin S + y_0 \end{cases} \quad (7)$$

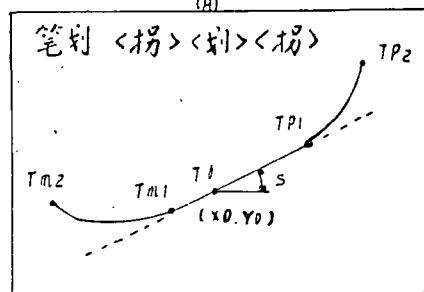
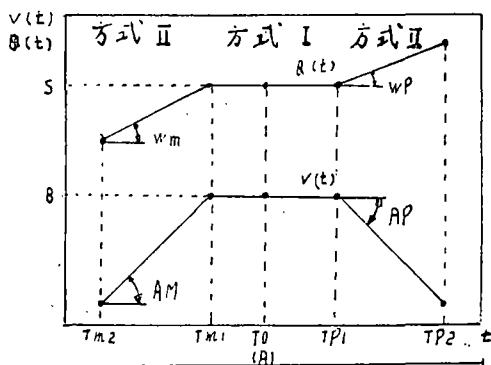
<拐>: 能满足式 (8) (9) (10) 的曲线笔点运动

$$\dot{V}(t) = V(t) \cdot \exp i(t) = (At + B) \cdot \exp i(\omega t + s), (\omega \neq 0) \quad (8)$$

$$\begin{cases} x(t) = \frac{1}{\omega} (At + B) \cdot \sin(\omega t + s) + \frac{A}{\omega^2} \cos(\omega t + s) + \frac{B}{\omega} \sin S - \frac{A}{\omega^2} \cos S + X_0 \\ y(t) = \frac{1}{\omega} (At + B) \cdot \cos(\omega t + s) + \frac{A}{\omega^2} \sin(\omega t + s) + \frac{B}{\omega} \cos S - \frac{A}{\omega^2} \sin S + Y_0 \end{cases} \quad (9)$$

$$\begin{cases} x(t) = \frac{1}{\omega} (At + B) \cdot \cos(\omega t + s) + \frac{A}{\omega^2} \sin(\omega t + s) + \frac{B}{\omega} \cos S - \frac{A}{\omega^2} \sin S + Y_0 \\ y(t) = \frac{1}{\omega} (At + B) \cdot \sin(\omega t + s) + \frac{A}{\omega^2} \cos(\omega t + s) + \frac{B}{\omega} \sin S - \frac{A}{\omega^2} \cos S + X_0 \end{cases} \quad (10)$$

式 (4) 中算子“|”和“。”表示算子两侧的笔划元素要求连续连接的关系为逻辑“或”和逻辑“与”。图 7(A) 表明在<拐>·<划>·<拐>类型的笔划综合情况下, 参数 A、B、W 和 S 的物理意义。该图中, 方式 I 和方式 II 分别表明<划>和<拐>的综合区间。图 7(B) 表明由图 7(A) 中的参数综合出的<拐>·<划>·<拐>类型的笔划。图 8 示出应用简化后的笔点运动模型综合出的五种典型的笔划。



(B)

图 7

笔划元素	笔	划
<点>	•	•
<划>	—	—
<拐>	⌒	⌒
<拐>·<划>·<拐>₁	⌒—⌒	⌒—⌒
<拐>·<划>·<拐>₂	⌒—⌒	⌒—⌒

图 8

图9以框图示出了笔划提取的操作程序。我们从线性笔划线段表中选择最长的线性笔划线段,同时,对笔点运动模型给出此笔划线段的位置 (x_0, y_0) 以及方向 (S) 信息作为笔划综合时的初始条件。然后应用这些初始条件和书写周期 $T = t - t_m$ 的模型合成出线性笔划。在这里,参数 B 保持一恒定值(平均笔点速度)。

综合出的笔划与由参考图形表给出的标准化输入图形之间的距离 D 由比较器测量。距离 D 由式(11)确定:

$$D = \frac{1}{N} \sum_{n=1}^N \left\{ \min_{r \in R} \left[\sqrt{(X_{sn} - X_{or})^2 + (Y_{sn} - Y_{or})^2} + P_1 \cdot d_1(\theta_{sn}, \theta_{or}) \right] + P_2 \cdot d_2 \right\} \quad (11)$$

式中, N 是综合出的笔划长度; $X_{sn}, Y_{sn}, \theta_{sn}$ 是综合出的笔划上的第 n 元素的位置和方向; $X_{or}, Y_{or}, \theta_{or}$ 是标准化输入图形中第 r 个黑元素的位置及它与相邻的八个元素连接的方向。 θ_{or} 可能取两个以上的方向值。 $(\theta_{or} = \theta_{or}^1, \theta_{or}^2, \dots, \theta_{or}^k, \dots, \theta_{or}^8)$ 。 R 是由两个 90° 的扇形组成的集,每一个扇形都与 θ_{sn} 正交。 P_1 和 P_2 是惩罚系数, d_1 和 d_2 是由式(12)和(13)确定的函数:

$$\begin{cases} d_1(\theta_{sn}, \theta_{or}) = 0 & (\min_k |\theta_{sn} - \theta_{or}^k| < \theta_r) \\ d_1(\theta_{sn}, \theta_{or}) = 1 & (\min_k |\theta_{sn} - \theta_{or}^k| > \theta_r) \end{cases} \quad (12)$$

$$\begin{cases} d_2 = 0 & (R(X_{or}, Y_{or}) \neq \phi) \\ d_2 = 1 & (R(X_{or}, Y_{or}) = \phi) \end{cases} \quad (13)$$

式中, $R(X_{or}, Y_{or})$ 是属于区域 R 的 (X_{or}, Y_{or}) 的集; θ_r 是阈值。图10是这些定义的图示说明。

误差估算级根据是将笔划提取结束还是需要再次进行综合的 D 的值来作出判定。若判定需要再次进行综合,则参数控制级根据参数修正规则来修正模型中的每一个参数值,并对笔划综合级给出再次综合的命令。当综合出的线性笔划与参考图形笔划足够匹配时,则参数控制级向笔划综合级给出一个稍长的书写期间 $(t_p + \Delta t)$ 与/或 $t_m - \Delta t$,而这就综合出稍长的笔划。

一旦 D 值超过某一阈值 D_r ,参数 X_0, Y_0 以及 S 就被修正,并综合一个新的笔划。但如果再次综合也未成功地将 D 值减小到比 D_r 值小的情况下,则线性笔划提取便告结束。然后,用类似的手续来综合与线性笔划两端连续地相接的曲线笔划。与线性笔划综合的情况一样,参数 A 和 W 的

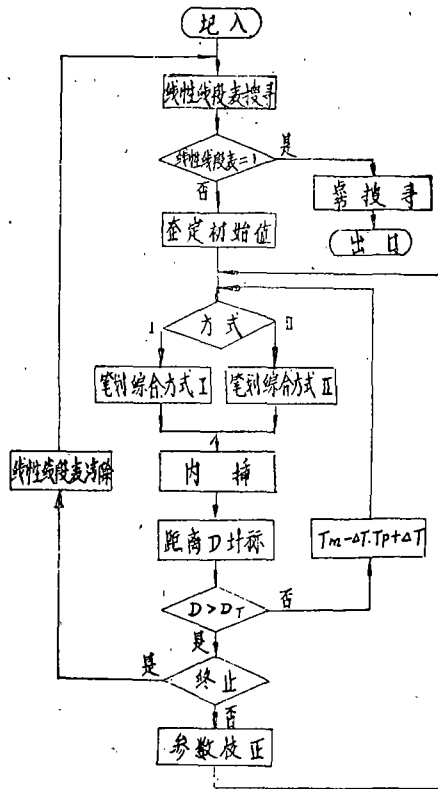


图9

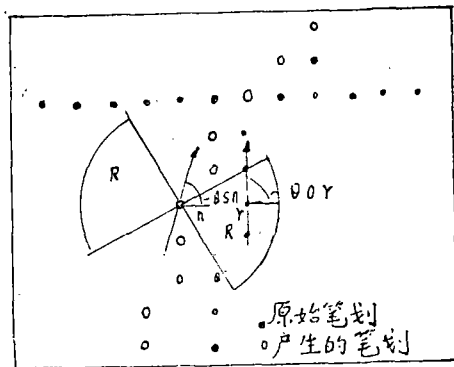


图10

值也逐渐地修正。为了获得最小的收敛时间,运用了试探法和最陡斜坡法作为参数修正规则。

将笔划综合的最后结果看作为所提取的笔划,将该笔划的类型、长度以及位置存储在一个提取出的笔划表中。然后,将它运用到下一个汉字识别过程。线性线段表中与综合出的笔划位置和方向最相近的线性笔划线段即予勾除。从表的其余部分选出最长的线性笔划线段,逐次进行笔划提取直到该表中的所有条目均被勾除为止。

笔划分类级将所抽取的每一笔划转换成13位的笔划码(其中3位表示笔划类型,2位表示笔划长度,4位表示X的位置,4位表示Y的位置),如图12所示。图13示出了笔划码中所用的7种笔划类型。

文字识别级和笔划字典组成了接受笔划码链的自动装置。图14表明笔划字典的结构。该自动装置首先考查所提取的笔划的序号;然后检验是否存在典型的笔划码。如果需要,也考查笔划之间的点和互连关系。当码链能被接受时,该自动装置发出一相应于被识汉字的代码信号,如果码链不能被接受时,便发出一个拒识信号。

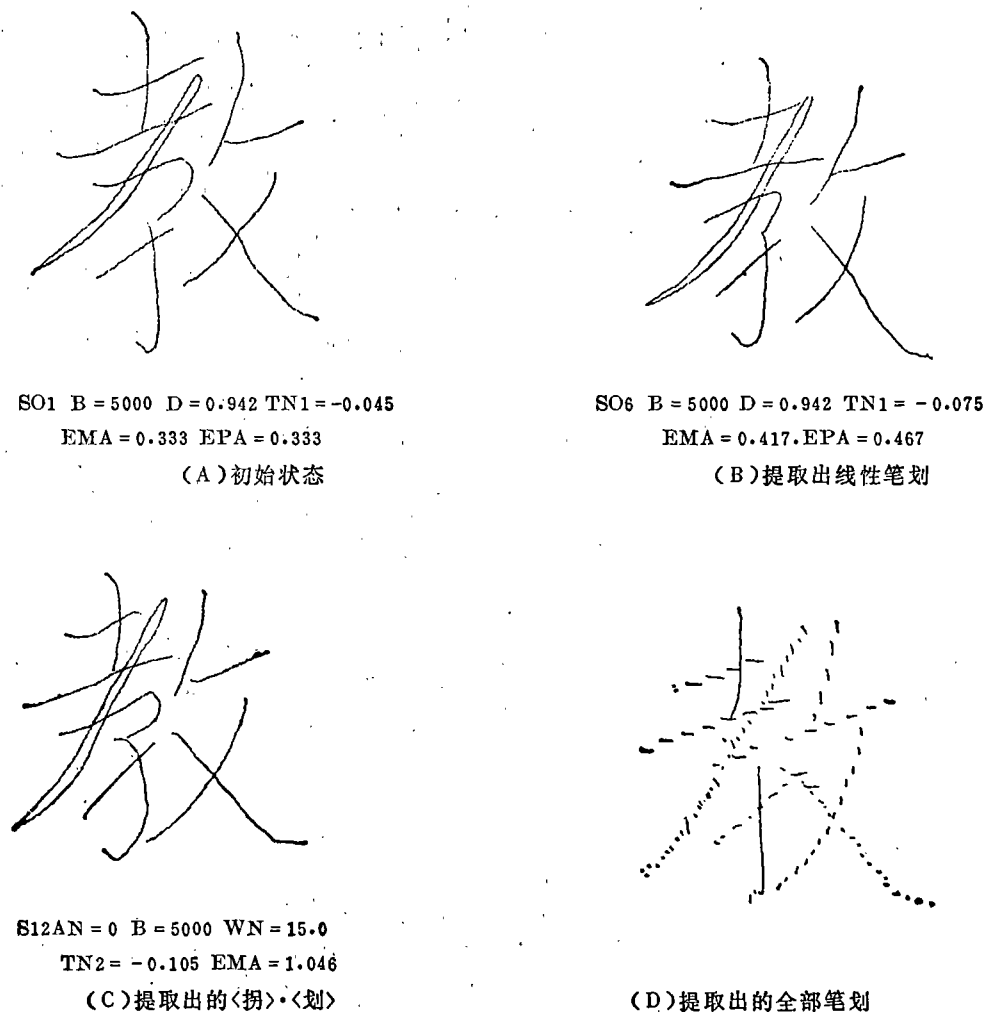


图 11

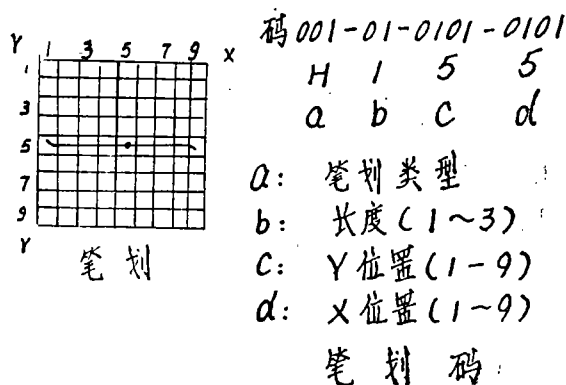


图 12

序号	码	笔划
1	H	—
2	V	
3	R	/
4	L	\
5	SR]
6	SL	[
7	D	

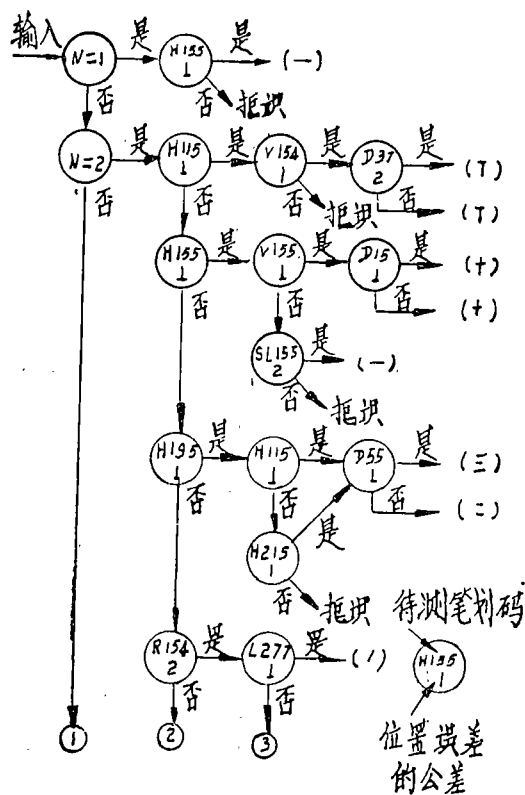


图 13

图 14

识别试验 应用PDP-9型计算机,对识别方法进行了模拟试验,全部模拟程序都用图形处理汇编语言⁽¹²⁾来书写。使用了500个简单的手写体汉字来建立笔划字典(从一般教学汉字中选出的100个汉字由5名大学生写出),另取500个手写体汉字进行识别试验。最后的结果是正确识别率82.4%。如表Ⅲ所示,识别所需的平均计算时间为 $12 + 8N$ 秒(N 是提取的笔划数)。

表 I

识别率82.4%

误识率 7.4% (不妥当的笔划提取1.2%不良的笔划字典6.2%)

拒识率10.2% (不妥当的笔划提取1.4%不良的笔划字典8.8%)

结论 尽管试验所用汉字的数量不多,但所介绍的方法还是令人满意的。

用分析综合方法进行笔划提取过程,在97%以上的案例中,均可精确地工作。这种结果证明手写体汉字按照笔划线段法是相当可靠的。另外,有待今后解决的一个问题是减小计算时间;另一个问题是,为了增加本方法识别汉字的数量,必须能自动地编辑笔划字典。

汉字还有另一个特点即局部图形分类结构。大部分汉字都可很容易地分解成诸如偏旁、字根、本体、冠盖等七类局部图形。借助汉字这个特点,将有助于今后提出简洁的笔划字典。

参 考 文 献

- [1] Casey R. and Nagi G., Recognition of Printed Chinese Characters, IEEE Trans., Vol. EC-15 (1966):1(Feb.), 91.
- [2] Nakano Y. and Nakata K., Classification and Recognition of Chinese Characters by Marginal distributions, Automaton Research Group Report of Japanese Institute of Electronic Engineers Vol. A71. (1971):52(Sept.)
- [3] Yoshida M., A C Type Voltage Pen System—A New Graphic Input System, Electronic Computer Research Group Report of Japanese Institute of Electronic Engineers, vol. EC71 (1972):53(Jan.)
- [4] Eden M., and Halle M., The Rharacterization of Cursive Writing. Symposium on Information Theory, London, 1962.
- [5] Edon. M. Handwriting and Pattern Recognition. IRE Trans. Vol. IT-8 (1962) :2(Feb.), 160
- [6] Mermelstein P. and Eden M., Experiments on Computer Recognition of Connected Handwritten Words, Information and Control, vol. 7(1964):2(June), 255.
- [7] Van der Gon D. and Turing J., A Handwriting Simulator, Physics in Medicine and Biology. vol. 6 (1962):3(Jan.), 407.
- [8] MacDonald J., Experimental Studies of Handwriting Signals, M.I.T. Technical Report*443, 1966
- [9] Yoshida M. and Ogata Y., "Handwriting Analysis And Synthesis. Journal of The Radio Research Ladoratories. (1966):65(Jan.)
- [10] Stevens K. N. "Toward a model for. Speech Recognition. "JASA, Vol. 32. *I. P47, January (1960); I(Jan), 47
- [11] Halle M. and Stevens K.N., Speech Recognition, A Model and a Program for Research. "IRE Tra ns., Vol. IT-8 (1962):2(Feb.)
- [12] Yoshida M., A Pattern Processing Compiler, Review of The Radio Research Laboratories, Vol. 14*, (1968):74(Sept.), 505.