第 14 卷 第 3 期
1988 年 5 月

自 动 化 学 报
ACTA AUTOMATICA SINICA

Vol. 14, No. 3
May, 1988

在线手写汉字识别的字形结构排序法

刘迎健　　戴汝为

（中国科学院自动化研究所）

摘　　要

汉字是二维平面上的线划图形，在线汉字识别的一个有利条件是利用书写时的笔段顺序信息，从而采用一维的表示．然而不同的人书写同一个字时笔顺会有所不同，这就给汉字的处理与识别带来困难．本文给出一种以笔段为基础，仅依赖汉字字形结构的排序方法，把二维空间的笔段在一维空间排出稳定的次序．这一次序与笔划的书写序无关．这就为在线手写汉字的识别打下了良好的基础．

一、汉字识别的结构层次划分

汉字从组成结构上可以分成 4 级：(1)整字级，(2)部件级(边旁、部首、字根)，(3)笔划级(落笔与抬笔之间笔尖的轨迹)，(4)笔段级(笔划中方向不变的最大线段)．其中整字 7000—10000 种，部件 200—500 种，笔划 40—80 种，相邻级别的种数相差约一个数量级，而且级别越低结构越简单．以上结构分层是很自然的．　在向别人解释一个字的写法时，常常是一级级解释下去，直到对方明白为止．　用句法方法[1]识别汉字时，也要对汉字的结构进行划分．　这样可以逐级用一小组简单的基元和文法规则来描述一大组复杂的模式．目前国内外有关汉字的在线识别方案中，一般仅分出两层：笔划级与整字级，一般均以笔划作为基元，以笔划书写次序(即笔顺)作为排序方法形成一维空间的笔划链，进行匹配运算．按理说大部分汉字是存在规范笔顺的，但由于每个人识字的年代不同，地区不同，老师不同，具体环境不同等因素，使人们对每个汉字的笔顺习惯不一定符合规范．许多字根就存在几种常见的笔顺，甚至说不清楚哪种是规范的．　为了解决不同笔顺的汉字识别问题，一般有以下两种途径：

（1）建立标准的识别字典(包括匹配时的标准序列和句法中的规则)时应尽量考虑到每一种书写笔顺．

（2）在匹配运算中使用句法误差分析方法，尽量校正笔顺误差．

第一种方法增加存贮的开销，第二种方法增加时间的开销．　在设计一个误识率和拒识率小于某一指标的实用汉字在线识别系统时，就会发现这种开销大到难以实现，即存贮占用过大或查找时间太长，这一问题就成为在线识别长期没有达到实用化的基本原因之一．

本文于 1987 年 1 月 7 日收到．

本文给出一种以笔段为基础，仅依赖汉字字形结构的排序方法，即二维空间的笔段可以在一维空间排出稳定次序．这一次序与笔划的书写顺序无关，这就为在线手写汉字识别打下了良好基础．

二、汉字笔段间的位置关系

在文章[2]中，笔者提出了一种识别笔段的模糊属性自动机，所有楷书的笔划可以通过模糊属性自动机分解成十一种广义笔段，见图 1． 所有楷书汉字都由这十一种广义笔

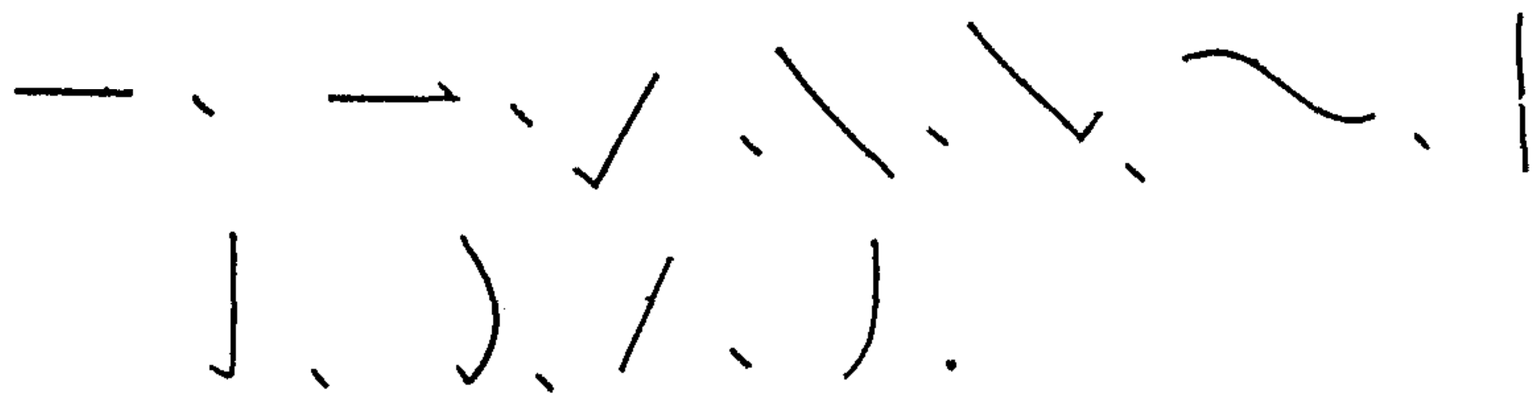


图 1

段以不同长短和不同结构组成．这十一种广义笔段中又以一、丨、丿、乀 四种最为基本，可以作为基本元素(不可再分的基元)，其它的广义笔段可以用这四种去近似，如图 2 所示．

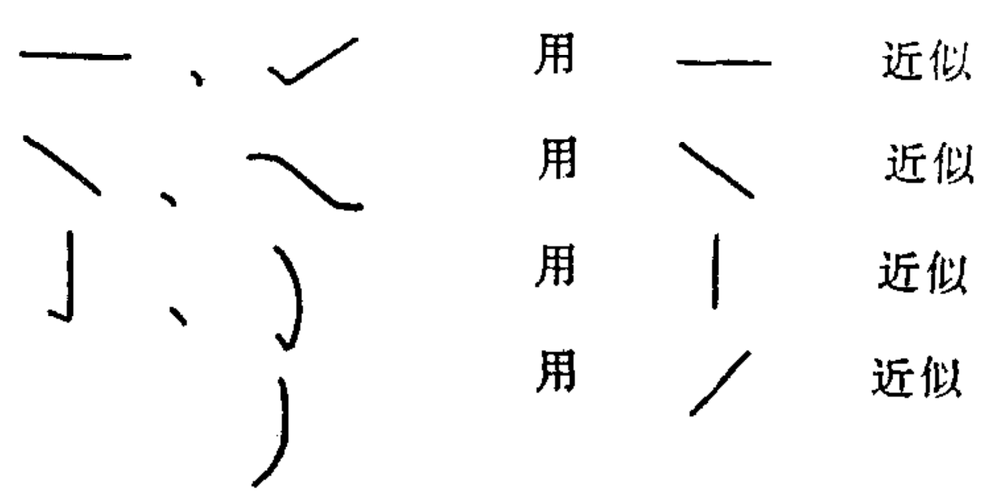


图 2

近似后，笔段间关系简单明确，便于分析，且大多数汉字用这四种笔段近似后，仍然可以正确辨认．

下面以四种笔段为代表，分析它们在组成汉字时相互之间的位置关系，这些关系对这 11 种广义的笔段也是成立的．

首先给出以下几点解释．

(1) 标号点．对基本线段，在三个点上给予标号，起笔点与笔段的终止点分别标以

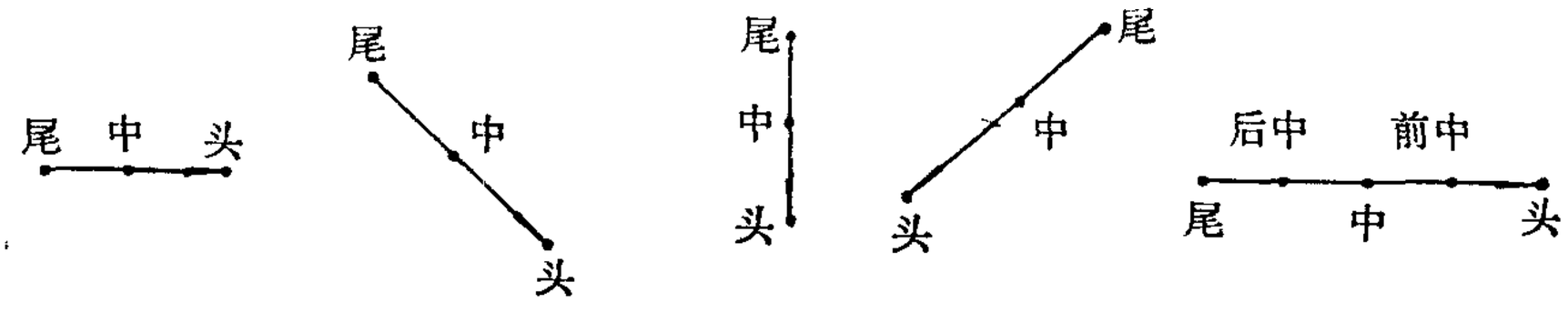


图 3

“尾”与“头”，并从尾到头用向量表示，尾头之间的中点标以“中”．在需要对位置关系精确描述时，还用到以下两个辅助标号点：

前中：基本笔段头与中之间的中点，

后中：基本笔段尾与中之间的中点．（见图 3）

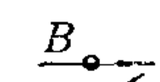
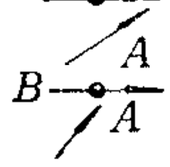
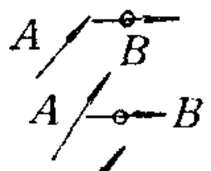
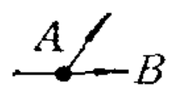
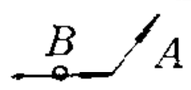
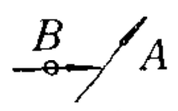
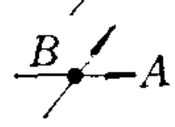
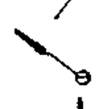
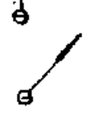
（2）两笔段的连接．对于两笔段 A 与 B，如果能在 A 上找到一点 C_a，在 B 上找到一点 C_b，使 C_a 到 C_b 的距离不大于给定常数 e，则称 A 与 B 连接，其中 e 是大于或等于 0 的实数．

（3）两笔段分离．如果两笔段 A 与 B 不连接，即 A 上任何点到 B 上任何点之间的距离大于给定常数 e，则称 A 与 B 分离．

（4）两连接笔段之间的关系．如果两笔段 A 与 B 连接，设 a 与 b 是两笔段间的最近点，$a\in A$，$b\in B$，a 到 b 的距离为 d，$0\leqslant d\leqslant e$，从 a 向 A 笔段上三个(或五个)标号点(头、中、尾)测距离，记下距离最短的那个点的标号 P_a．从 b 向 B 笔段上三个(或五个)标号点(头、中、尾)测距离，记下距离最短的那个点的标号 P_b．称 A 对 B 是 P_aP_b 关系．

（5）两分离笔段之间的关系．如果 A 与 B 是两个分离的笔段，分别把 A、B 向 X、Y 轴上投影．在 X 轴上的投影为 x_a、x_b，在 Y 轴上的投影为 y_a、y_b．取 $x=\max[x_a, x_b]$，$y=\max[y_a, y_b]$．x 的两个端点把 X 轴分成三段，即左段、中段、右段．判定相对笔段

表 1　两笔段之间的 17 种关系

序　号	关系名称	关系表示	对偶关系
1	(头头)(hh)		(头头)(hh)
2	(头中)(hm)		(中头)(mh)
3	(头尾)(ht)		(尾头)(th)
4	(中尾)(mt)		(尾中)(tm)
5	(尾尾)(tt)		(尾尾)(tt)
6	(尾中)(tm)		(中尾)(mt)
7	(尾头)(th)		(头尾)(ht)
8	(中头)(mh)		(头中)(hm)
9	(中中)(mm)		(中中)(mm)
10	(左上)(la)		(右下)(rd)
11	(中上)(ma)		(中下)(md)
12	(右上)(ra)		(左下)(ld)
13	(右中)(rm)		(左中)(lm)
14	(右下)(rd)		(左上)(la)
15	(中下)(md)		(中上)(ma)
16	(左下)(ld)		(右上)(ra)
17	(左中)(lm)		(右中)(rm)

(A相对于 B，B 相对于 A) 的中点的坐标落在哪一段. 如考虑 A 相对于 B，而 $\max[x_a, x_b] = x_a$ 的情况，则分析 B 的中点在 x 轴上的哪一段，记为 $\bar{Q}_a$；再把 $\bar{Q}_a$ 取反，即右 → 左、左 → 右、中 → 中，则得到 A 对于 B 的段号 Q_a，也可用类似办法判定相对笔段的中点在 Y 轴的哪一段，记下段号 Q_b，称 A 对 B 为 Q_aQ_b 关系.

实际上，如果相对于 B 而言，即以 B 为参考，可分成八个方向，即左上 (la)、中上 (ma)、右上 (ra)、右中 (rm)、右下 (rd)、中下 (md)、左下 (ld)、左中 (lm).

(6) 设 A 与 B 的关系为 r，即 ArB，B 对 A 的关系为 r'，即 $Br'A$，则称 r 和 r' 互为对偶关系.

归纳起来，组成汉字的基本笔段间的相互位置关系可分成两大类：连接与分离，笔段间的交叉关系可看成连接，即中点和中点的连接. 连接关系有九种，分离关系有八种. 表 1 是两大类17种关系的名称，相应的英文字母标号、关系的表示以及对偶关系.

三、汉字的结构与笔段顺序

在汉字的部件中，笔段之间的相互位置是比较稳定的，部件之间的笔段位置关系有点模糊. 实际上我们可以根据汉字本身的结构，把组成该汉字的笔段加以排序. 一般说来，不可能像建立模板那样，对每个汉字建立一个笔段顺序的模板，而是建立有关笔段之间顺序的一种标准. 比如任意两笔段 T_i 与 T_j，它们在位置上有某种关系，则可以预先确定哪一笔段应"优先"于哪一笔段. 如果两个笔段出现的顺序要求满足预先的规定，则给予一定的标记，例如用 1 标志. 如果出现的顺序不满足预先规定，则用 0 标志. 为说明起见，给出两种笔段分离情况下的一个关系表，如图 4 所示.

A	关系	B	值	A	关系	B	值
X	左上	X	1	X	右下	X	0
X	中上	X	1	X	中下	X	0
X	右上	X	1	X	左下	X	0
X	右中	X	0	X	左中	X	1

图 4

另外图 5 给出仅包括 4 种基本笔段，且为中中关系时汉字结构关系表.

对于一个由 n 段笔段组成的汉字，在识别过程中，由模糊属性自动机不断分解出广义笔段，并给出近似的基本笔段序列 T_1，T_2，$\cdots T_n$. 可以根据这 n 个笔段之间的关系，来定义一个方位矩阵 Y. Y 的第 i 行第 j 列 ($i < j$) 的元素用笔段 T_i 相对于笔段 T_j 的关系 t_{ij} 表示，第 j 行第 i 列的元素则用 t_{ij} 的对偶关系表示，而对角线上的元素 t_{ii}，$i = 1, 2, \cdots n$ 用 ϕ 表示. 这个矩阵 Y 充分反映了 n 个笔段相互间的位置信息，Y 如图 6 所示.

任意两笔段之间，例如第 i 段 T_i 与第 j 段 T_j 之间位置关系为 t_{ij}，如果按照汉字本身的结构，T_i 应优先于 T_j，则令 $t_{ij} = 1$，$t_{ji} = 0$. 当书写时出现的顺序与结构关系不符，则令 $t_{ij} = 0$，$t_{ji} = 1$. 以 0 或 1 作为元素的方位矩阵，对分析各笔段间的相互位置关系很有帮助. 书写汉字时，每当出现一个新的笔段 T_i，就要分析与已经出现了的 $(i-1)$ 个笔段之间的相互位置，测出 T_i 与 T_1、T_2、$\cdots T_{i-1}$ 之间的位置关系. 如果 T_1、T_2、$\cdots T_i$ 出现的次序符合结构关系表中的规定，那么就可以得到一个 $i \times i$ 的矩阵. 这个矩阵的

（中中）　关系				
B / A	—	＼	|	／
—	0	1	1	1
＼	0	0	0	0
|	0	1	0	0
／	0	1	1	0

图　5

ϕ				
t_{21}	ϕ			
t_{31}	t_{32}			
⋮	⋮	⋱		
t_{n1}	t_{n2}		t_{nn-1}	ϕ

图　6

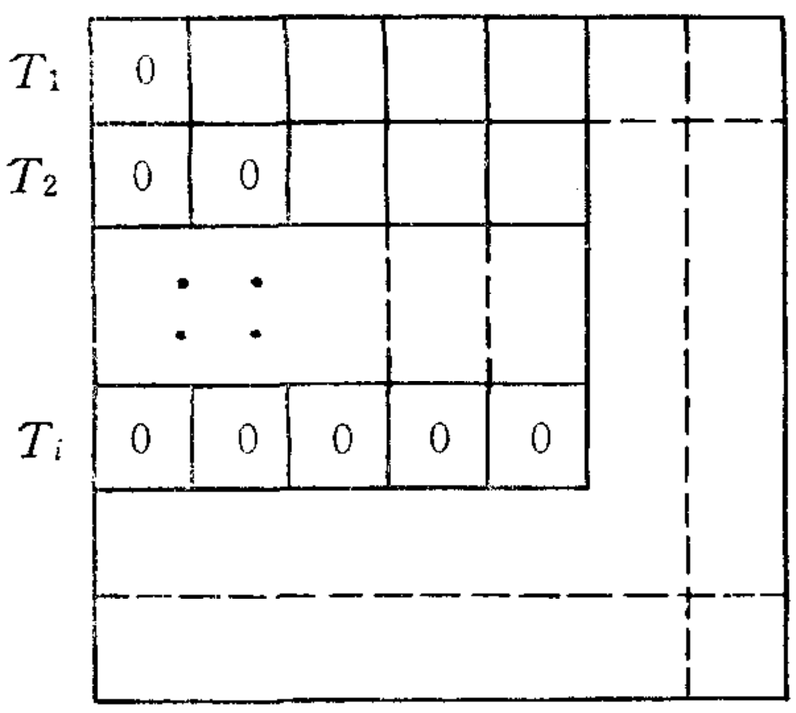


图　7

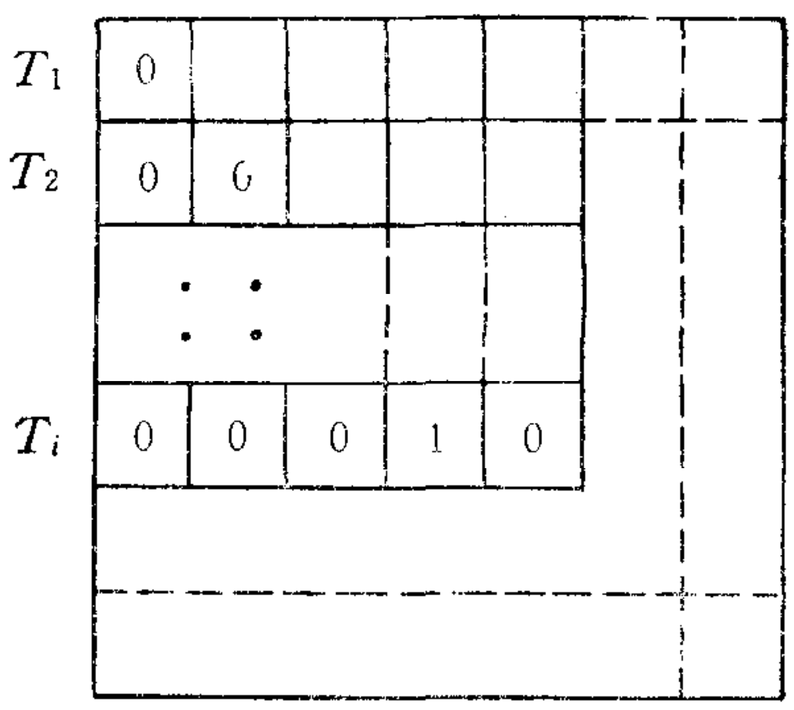


图　8

下三角(令Y矩阵的对角线上元素 $\phi = 0$) 上的元素全为0，上三角上的元素全为1，如图 7 所示. 如果 $T_1, T_2, \cdots T_i$ 出现的次序与汉字结构关系表中的关系不符，那么 $i \times i$ 矩阵的下三角上的元素就不全为 0. 例如 $T_1, T_2, \cdots T_{i-1}$ 的先后次序符合结构关系，而 T_i 与 T_{i-1} 的次序不符合结构关系，那么在 $i \times i$ 矩阵的第 i 行中就有 $t_{i,i-1} = 1$，而上三角中就有 $t_{i-1,i} = 0$，如图 8 所示.

总而言之，一个由 n 段笔段组成的汉字，如果笔段出现顺序按照结构的优先关系出现，那么就可得一下三角为0，上三角为 1 的 $n \times n$ 矩阵. 这种情况下，第 i 行的元素总和为 $(n - i)$，$i = 1, \cdots n$，即从大到小. 在这一想法的启迪下，给出下述把不同笔顺序列化为按结构排序的方法.

四、汉字笔段结构排序法

为叙述简单起见，考虑由四种基本笔段组成汉字. 用 $V = \{—、／、|、＼\}$ 表示基本笔段集合，$R = \{(h, h), (h, m), \cdots (l, m)\}$ 表示笔段间 17 种关系集合 $X = \{0, 1\}$. 通过以下的步骤对 n 个笔段排序.

(1) 对方位矩阵Y中的每个元素 t_{ij} 与相应的笔段 T_i 与 T_j 做W映射，$W(T_i, T_j, t_{ij}) = Z_{ij}$，形成一个新的 $n \times n$ 矩阵 Z. Z 中的元素为 0 或 1，反映了 T_i 对 T_j 的位置关系是否与汉字的结构关系相符合.

(2) 设 X 是 $n \times 1$ 列矩阵,其中的元素都为 1,即 $X' = (1,1,\cdots 1)$. 做 Z 和 X 的矩阵乘积 $P = Z \times X$, 则 P 中的元素为 $p_i = \sum_{j=1}^{n} Z_{ij}$.

(3) 用 P' 表示 P 的转置,则 P' 是元素为正整数及 0 的 $1 \times n$ 行矩阵,把 P' 中各元素 $p_1, \cdots p_n$ 的值赋给对应笔段序列 T 中各元素 $T_1, \cdots T_n$, 这样各笔段就可以根据所赋的数值相互比较大小,并按从大到小重新排序. 当碰到相同值元素时,再做辅助映射 V. V 是 T 的子集到 T 的子集的映射,通过映射使若干个相同值的元素之间也有一个确定的序. 排序时,每做一次笔段对换,方位矩阵也做相应的行列对换,最后得到新笔段序列 T' 和新方位矩阵 Y'. 把 T' 称为汉字结构序, Y' 为汉字结构矩阵.

以上方法的直观意义如下: 给定两个笔段 T_i、T_j 和它们之间的位置关系, 决定谁"优先". 当 T_i 优先于 T_j 时,表中对应次为 1,否则为 0. 然后把笔段 $T_i(i = 1, 2, \cdots n)$ 与其它笔段比较的结果相加,并按和的大小重新排序,这样就把对应于数值较大的笔段排在较前面. V 决定等值元素的序, V 可选择笔段中的 x、y 坐标的大小、距原点的距离等因素排序. 例. 下面以两种笔划顺序书写的"禾"字为例来说明排序的过程,如图 9 所示.

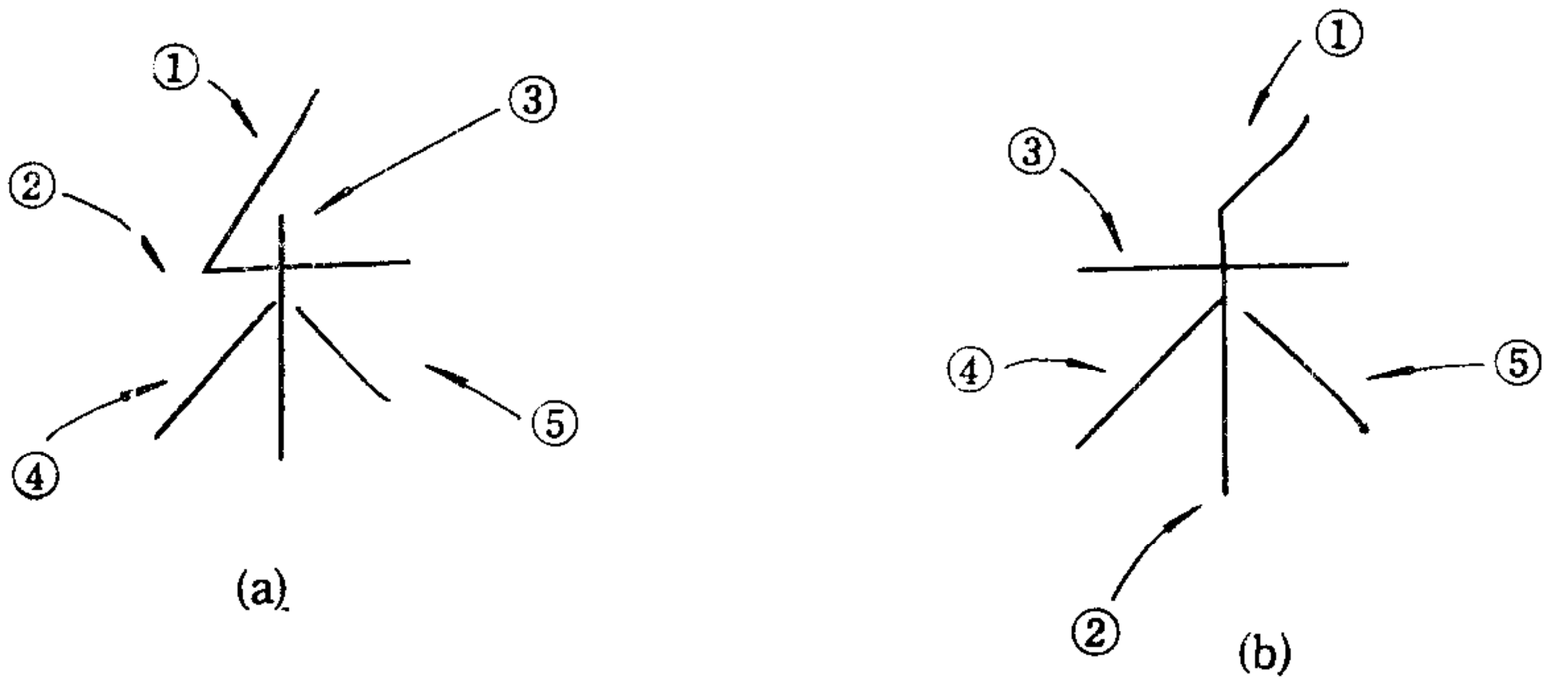


图 9

两种笔段序列 T_a 与 T_b 分别为:

$T_a =$ ／、一、丨、／、＼,

$T_b =$ ／、丨、一、／、＼.

两种位置关系矩阵 Y_a, Y_b 分别示于图 10.

$Y_a =$

ϕ	头尾	中尾	中上	左上
尾头	ϕ	交叉	中上	中上
尾中	交叉	ϕ	中尾	中尾
中下	中下	尾中	ϕ	左中
右下	中下	右中	右中	ϕ

$Y_b =$

ϕ	头尾	中上	中上	中上
尾头	ϕ	交叉	中尾	左中
中下	交叉	ϕ	中上	中上
中下	尾中	中下	ϕ	左中
中下	右中	中下	右中	ϕ

图 10

下面给出矩阵 Y_a、Y_b 经过 W 映射后所得的矩阵 Z_a、Z_b, 分别示于图 11。 求出

$Z_a =$

0	1	1	1	1
0	0	1	1	1
0	0	0	0	1
0	0	1	0	1
0	0	0	0	0

$Z_b =$

0	1	1	1	1
0	0	0	0	1
0	1	0	1	1
0	1	0	0	1
0	0	0	0	0

图　11

$$P_a = Z_a \times X \text{ 以及 } P_b = Z_b \times X,$$

$$P'_a = (4, 3, 1, 2, 0),$$

$$P'_b = (4, 1, 3, 2, 0).$$

根据 P'_a 中各元素的值的大小，把相对应的笔段进行排序，得到

$$T'_a = ／、一、／、丨、＼.$$

根据 P'_b 中各元素的值的大小进行排序，得到

$$T'_b = ／、一、／、丨、＼,$$

$T'_a = T'_b$，很明显也有 $Y'_a = Y'_b$.

上述按结构排序的方法，虽然不够严格，但关于在线手写汉字识别的大量实践工作表明它是非常有效的. 在汉字的部件中，笔段之间的相互位置是比较稳定的，在部件之间的笔段位置关系有点模糊，但映射成 0、1 之后也相当稳定.

五、笔段及字符的识别

根据上述排序的办法，就能够以笔段作为基本元素，用笔段及其出现的次序来描述二维平面上的汉字. 从原理上讲，汉字的识别与笔段的识别也就没有什么区别了. 这一点还可进一步加以说明. 一个汉字的笔段之间可能是分离的，但每个笔段标有尾、中、头三个点，按照顺序把笔段序列 $T_1, \cdots T_n$ 中相邻两段的中点连接起来，即 $T_i (i = 1, \cdots n-1)$ 与 T_{i+1} 相连，得到以 T_i 的中点为尾，T_{i+1} 的中点为头的向量 u_i. 再把 T_n 与 T_1 以同样方法连接，得到向量 u_n. 如果 T_i 与 T_{i+1} 二者为(中中)关系，则 $u_i = 0$，就可得到与汉字相对应的序列 $u_1, u_2, \cdots u_n$. 这一序列相邻两段是连接的，u_i 可以用如下←、→、↑、↓、↗、↙、↖、↘八个向量近似表达. 一个由 n 笔段组成的汉字，就对应于由八个基元组成的串 $u_1, u_2, \cdots u_n$，这样的符号串类似于笔段的结构，所以可以利用能有效地识别笔段的模糊属性自动机来进行识别.

六、结　束　语

在线手写汉字识别所遇到的问题之一是不同人书写同一汉字时笔划顺序不可能一致. 本文提供以笔段为基础，按结构关系排序的方法，具有稳定的笔段顺序再加上另一文中提供了一种模糊属性自动机，能有效地识别笔段. 通过这两项工作，把在线手写汉字识

别工作推向实用化。以下是在 PC—XT 上试验的结果。硬件条件是：时钟 4.77MHz，存贮容量 64kB，识别时间小于 0.4 秒。对书写者要求是书写较工整。经几分钟示范及试用后开始记录识别率，见表 2。

表 2

结构划分方式	初次使用识别率	经常使用后识别率	识别范围
笔划一整字	52%	95%	4000
笔划一笔段一整字	70%	98%	7000
笔划一笔段一部首一整字	80%	99%	8000

参 考 文 献

[1] 傅京孙，模式识别及其应用，科学出版社，1983 年.
[2] 刘迎健，戴汝为，识别在线手写汉字的模糊属性自动机，自动化学报，**14**(1988)，2.

A METHOD OF LINESEGMENT ORDERING BY CHARACTER STRUCTURE FOR ONLINE HANDWRITTEN CHINESE CHARACTER RECOGNITION

LIU YINGJIAN DAI RUWEI
(*Institute of Automation, Academia Sinica*)

ABSTRACT

A Chinese character is a 2-D line-drawn pattern. In order to get a 1 D representation, the order of strokes is often applied to online Chinese character recognition. But the order of strokes may be different for different writers. A method of linesegment ordering is proposed in this paper. It can provide stable order independent of the writing order of characters. On the basis of this method, high performance online handwritten Chinese recognition can be obtained.