第 24 卷 第 4 期
1998 年 7 月

自 动 化 学 报
ACTA AUTOMATICA SINICA

Vol. 24, No. 4
Jul., 1998

一种活动模板子结构引导的联机手写汉字识别方法[1)]

肖旭红 戴汝为

（中国科学院自动化研究所 北京 100080）

摘 要 结构匹配是一种有效的联机手写汉写识别方法，为了减少匹配运算，人们一直在寻求利用部分匹配的结果来引导整体匹配的方法. 在特征匹配与结构匹配综合的基础上，从 3.755 个一级国标汉字中提取出 45 个子结构，利用它们来引导结构匹配. 由于这些子结构总出现在字首或字尾，因而对它们的检测比较容易. 同时，通过建立子结构活动模板及设计子结构动态抽取算法，使得子结构匹配的准确度得到很大提高. 实验结构表明，该方法使结构匹配的运算量减少约 50%，并对类似的物体识别问题有一定的启发意义.

关键词 联机手写汉字识别，结构匹配，子结构，活动模板.

1 引言

随着便携式笔记本电脑的发展及联机输入接口性能的提高，联机手写汉字识别作为一种快捷的即写即识的汉字自动录入方式，再一次引起了广大研究者的兴趣. 在对规整楷书的识别取得了一定成绩之后，人们把研究的重点转移到笔顺变化及具有连笔的自由手写汉字的识别. 虽然有的识别系统较好地解决了带有连笔的汉字的识别[1]，而另一些系统较好地解除了对书字笔顺的限制[2]，但几乎还没有一个系统能同时解决这两个问题.

联机手写汉字一般都以笔划序列的方式存储，由于这一得天独厚的条件，基于笔划及笔段串匹配的结构匹配算法在联机手写汉字识别中得到了广泛的应用，通过对各结构基元赋予拓朴及几何属性，结构方法能在一定程度上吸收汉字的形变，但其运算复杂度较高，对变笔顺汉字进行识别尤其困难. 因此近年来有人趋向于用脱机汉字识别的方法来识别联机汉字，即将联机汉字转化为二维图象，利用特征匹配方法来识别联机汉字. 特征匹配方法的优点是模板及匹配算法的实现很容易. 由于在不存在联笔的情况下，笔顺的变化并不影响汉字的图象，因此用特征匹配方法识别没有连笔的联机手写汉字，可有效地消除笔顺对识别结果的影响，但当存在连笔时，两笔划间相连接的笔段导致汉字图象的变化，从而给识别结果带来影响，同时，目前的特征匹配算法对汉字形变的吸收能力相当有限，而由于人们对联机书写不习惯等原因，联机汉字的形变比脱机汉字大得多，因而单纯用特征匹配的方法进行识别是不够的. 作为一个典型的模式识别问题，联机识别技术的发展有赖于对汉字结构的描述及计算机对人脑识字过程模拟水平的提高.

1）国家自然科学基金资助项目.

收稿日期 1996-09-18

由于书写习惯等原因，在很多汉字中都存在这样的字根或子结构，绝大多数人总是在最初或最后的几个笔划中书写它们. 比如，当我们书写汉字“啊”时，总是先写偏旁“口”，再书写其它部分；而对于汉字“利”，总是先完成其它部分的书写，最后才写偏旁“刂”. 基于此，本文所介绍的系统利用特征匹配法对联机汉字进行粗分类，然后以字首或字尾的子结构引导属性串匹配来完成汉字识别的细分类过程. 由于这些子结构总出现在字首或字尾的笔划串中，因而只要有一个合适的匹配算法，对它们的检测是不困难的，当检测到待识别的汉字具有某一子结构时，只有包含该子结构的候选汉字才有资格与待识汉字进行结构匹配，因而大大地降低了结构匹配运算，同时也减少了连笔或笔顺变化导致的误识. 在本文的后续各节中，分别对识别系统模型、子结构的构造及匹配进行了介绍.

2 研究背景与系统模型

汉字识别的进展与人们对汉字结构的理解密切相关，而对汉字结构的理解又受到汉字的书写及人们对汉字的感知过程的影响，从而导致了一类与通常的汉字描述很类似的识别方法，即通过笔划间的关系识别出汉字部首或子结构，再通过这些已识别的子结构间的关系完成整个汉字的识别. 基于这种思想，早在八十年代末，刘迎建等就提出了一种启发式引导模板强制匹配的识别方法[3]. 这种方法从按上下序、左右序等不同顺序排列的笔段有序串中分离出字根并与相应的模板进行匹配. 但是什么时候从上下序的串中寻找字根，什么时候从左右序的串中寻找字根？这是一个相当困难的决策，几乎没有一个算法能对这一过程进行学习并做出正确的决策，即使我们通过手工操作，将每一个汉字的结构及搜索顺序“手把手”地教给计算机，由于手写汉字笔划数不固定及笔段分割算法不完善等原因，也无法确定到有序串的哪一个位置去寻找相应的字根，因而这种识别方法的实现是相当困难的.

为了充分地利用联机手写汉字的动态及静态信息，作者曾提出了一个利用图象特征匹配粗分类，结构匹配法细分类的联机手写汉字识别系统[4]，粗分类过程为每个待识汉字提供若干候选，细分类过程将待识汉字与其候选字的属性笔段串进行结构匹配，最后综合特征匹配与结构匹配的结果得到待识汉字的类别. 结构匹配是该识别系统的重要环节，导致其出现错误的原因主要有三个：首先，一个汉字由很多笔划组成，其中一个笔划的位置变化就有可能导致笔段的相对属性，诸如相对位置、各方向的笔划穿刺数等的变化，汉字笔划越多，造成相对属性不稳定的因素越多；其次，汉字笔划一旦过多，由于匹配方法的局限性，其结构的局部差异就有可能被众多的相似之处淹灭，不利于形似字，如“拔”与“拨”的区分；第三个因素是笔顺变化会在一定程度上影响结构的正确匹配. 结构匹配法的另一个重要缺陷是运算复杂度太大. 假如我们能根据汉字的部分结构来确定是否进行整字的结构匹配，就可大大地减少匹配运算；同时，把汉字分解成多个部件后，每一部件的结构相对简单，这也可以减少匹配错误.

在汉字笔划的起始或末尾存在很多结构稳定的部首或字根，由于它们位置的特殊性，比较容易把它们从笔划串中分离出来. 因此，本文定义了 45 类子结构，通过子结构检测来引导结构匹配. 整个系统框图如图 1 所示.

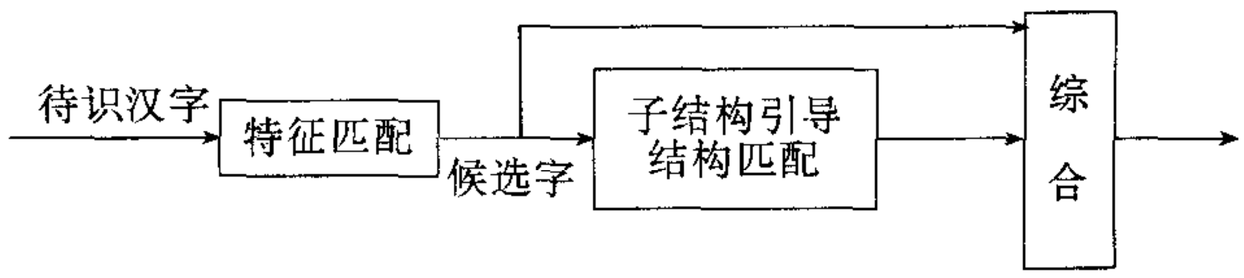

图1　联机汉字识别系统框图

3　活动模板子结构的构造

引导汉字识别的子结构的选择，主要遵循以下规则：

1)这些子结构出现在汉字的开始或末尾的笔划串中. 出现在汉字起始笔划中的子结构被称为字首子结构，出现在汉字终止笔划中的子结构被称为字尾子结构.

2)这些子结构必须具有较强的分类能力. 一个子结构的分类能力主要取决于两个因素：①是否容易与其它的子结构或笔划串混淆；②在候选字中出现的频数. 若初分类提供的待识汉字的所有候选字都包含某子结构，则它对这一汉字的匹配没有任何引导作用，因为所有的候选都必须与待识汉字进行匹配；另一方面，某一子结构在不包含它的待识汉字的候选字符集中出现越多，需要与待识汉字进行匹配的候选越少.

3)子结构的数目必须适中. 若所选择的子结构太少，则几乎所有候选都必须与待识汉字进行匹配；若所选择的子结构太多，则花在子结构检测上的时间增加且各子结构容易混淆，从而导致识别系统性能降低.

4)不同人对这些子结构的书写笔顺相对稳定，否则，由于匹配算法的局限，有可能导致子结构匹配失败，从而造成识别错误.

通过反复实验，作者挑选出 45 个子结构，它们在 3 755 个汉字中出现的总频数达 2 106. 表 1 中各项分别为这 45 个子结构的名称、序号及其在 3 755 个一级国标中出现的频数.

表 1　子结构名称、序号及出现频数

序号	名称	频数	序号	名称	频数	序号	名称	频数
0	刂	50	15	走	10	30	土	53
1	灬	22	16	马	24	31	山	25
2	页	30	17	雨	17	32	王	33
3	扌	215	18	车	25	33	户	7
4	氵	213	19	石	48	34	尸	32
5	木	133	20	proper	42	35	疒	44
6	竹	44	21	礻	25	36	广	35
7	禾	42	22	火	43	37	贝	23
8	钅	77	23	纟	75	38	巾	11
9	衤	20	24	目	26	39	又	14
10	饣	13	25	月	71	40	口	199
11	米	22	26	日	51	41	门	22
12	耳	13	27	女	52	42	讠	86
13	舟	12	28	犭	28	43	辶	44
14	虫	31	29	饣	20	44	穴	19

表中序号为 0—2 的子结构为字尾子结构，其它均为字首子结构. 从表中可以看出，所选择的子结构不包括“辶”、“亻”及“忄”等常见的偏旁，这主要是因为偏旁“辶”在字中的书写顺序很不稳定，有的人习惯先写偏旁“辶”，再写字的其它部件，而另一些人则正好相反；至于不选偏旁“亻”及“忄”，主要是因为这些偏旁容易与其它的部首及笔划串混淆.

3.1 汉字子结构属性的确定

子结构的序号是子结构的重要属性之一，据此，我们对每一个汉字赋予相应的属性. 当一个汉字包含某子结构时，该子结构的序号被作为该汉字的子结构属性，否则，该汉字的子结构属性被指定为-1. 若某汉字包含一个以上的子结构，则选择序号较小的子结构作为它的属性. 如汉字“热”既包含子结构“扌”又包含“灬”由于“灬”的序号小于“扌”，因此其属性被定为子结构“灬”的标号，即其序号为 1. 子结构序号的第二个作用就是反映了子结构的匹配顺序，若待识汉字的候选字集提供多个子结构，则从具有最小序号的子结构开始匹配. 这是因为在为子结构设计序号时，就考虑到了子结构间的部分包含、相似关系及一个汉字包含多个子结构的情况. 标号较小的子结构有可能包含标号大的子结构，如子结构“日”中包含子结构“口”. 当一个汉字包含多个子结构时，一般来说，其中标号较小的子结构具有较小的出现频数，因而分类能力较强，应该首先匹配.

3.2 子结构活动模板

基于子结构的识别方法主要须解决两个问题：一是子结构的抽取；二是对同一类型的子结构，当其处于不同字中的不同位置时，大小及笔划间的比例、位置关系也会相应地变化，从而给匹配识别带来困难. 为了解决第一个问题，本文方法除了采用字首子结构及字尾子结构外，还利用笔划数固定的正楷手写汉字建立模板及动态匹配规则来动态抽取待识汉字的子结构. 为了解决第二个问题，本文提出了活动模板的概念. 我们并不为每一个子结构建立由笔划串组成的固定模板，而是只记录组成每一子结构的笔划数目，匹配时再动态地从候选字的笔划串模板的起始或终止位置抽取相同笔划组成的笔划串作为该子结构对应于该字的活动模板，由于侯选字的笔划串模板是从笔划数固定的正楷学习得到的[4]，因而每一个子结构所包含的笔划数目也是固定的，从而可以轻易而准确地得到子结构的模板. 活动模板的建立，使我们从为每一个子结构建立很多个不同位置的模板的困境中摆脱出来，从而使利用子结构引导整字的匹配成为可能.

4 子结构引导的结构匹配

假设粗分类算法为待识汉字提供的第 i 个候选字用 $w_j(i=1,2,\cdots,M)$ 表示，若候选 w_j 中存在第三节中定义的子结构(即其子结构属性不为-1)，其子结构用 a_i 表示，则子结构引导的结构匹配过程如图 2 所示：从含有最小非负子结构属性的候选字开始(负的属性值表明该候选字不包含所定义的子结构)，从该候选的标准模板与待识汉字中抽取子结构并进行匹配，若匹配代价小于门限 th，则证明子结构检测成功，根据匹配结果动态地确定待识汉字中子结构的终止(对字首子结构)或起始(对子尾子结构)笔段的位置，从标准模板及待识汉字中删去子结构所包含的笔段，将所剩余的部件进行匹配；若子结构匹配不成功，即所有的子结构匹配代价 c_i 都大于 th，则将待识汉字与候选字的标准模板逐一进行匹配.

4.1　子结构的抽取与笔段排序

由于参考样本的模板具有固定数目的笔划，因而从其字首或字尾抽取出子结构是不成问题的，而待识样本则不同，由于可能存在的连笔，使得它所包含的子结构的数目小于或等于标准模板的笔划数，因此，子结构的抽取不再象参考模板那样直接，为此，本文方法对待识样本的子结构采用了动态的抽取策略.

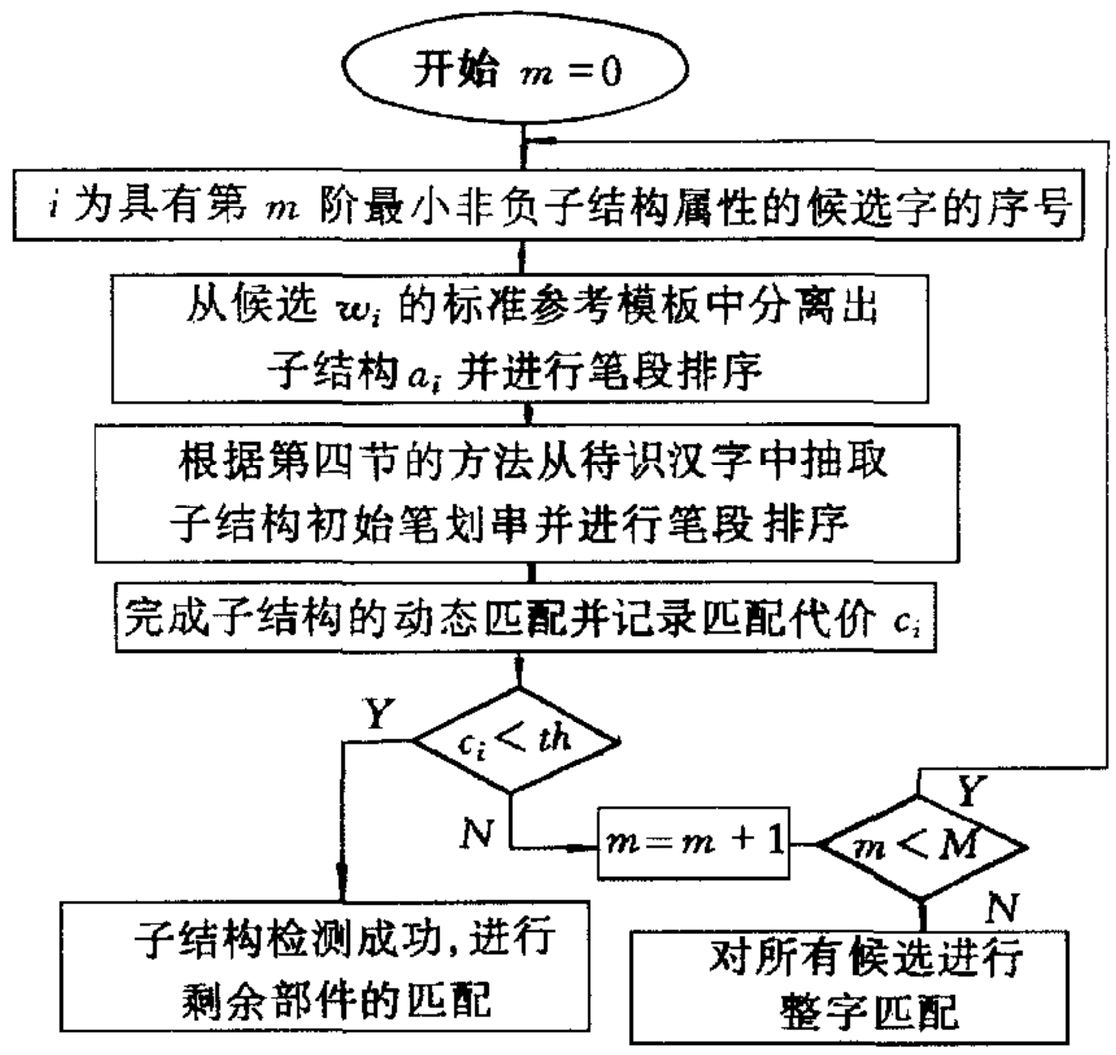

图 2　子结构引导的结构匹配过程

假设某待抽取子结构的标准参考模板由 n 个笔划组成，则首先以待识字的前 n 个笔划(对字首子结构)或最后的 n 个笔划(对字尾子结构)组成的笔划串作为该子结构的初始笔划串模板，然后将每一笔划分解成笔段[5]. 为了利于识别带有连笔的汉字，将每一笔划的终点与它相邻的下一笔划的起始点相连，这种两笔划间的连接线段被称为虚拟笔段. 各笔段顺序排列，形成初始笔段串. 虽然选择子结构时力求其笔顺稳定，但为了避免偶然发生的笔顺变化导致子结构检测错误，仍需对笔段重新排序，各笔段的先后顺序主要根据以下原则确定.

1)若两笔段间存在明显的上下或左右的关系，则这两笔段按从上至下或从左至右的顺序排列. 如图 3 中，笔段 l_1 与笔段 l_4 为左右关系，笔段 l_8 与笔段 l_6 为上下关系，因而笔段 l_1 应排于笔段 l_4 前面，笔段 l_8 位于笔段 l_6 之前.

2)若两笔段间不存在明显的上下、左右的关系，或为左下、右上关系，则这两笔段按书写的先后顺序排列. 如图 3 中笔段 l_4 与笔段 l_8 为右 * 关系(其中 * 表示关系不确定)，因而根据书写顺序，笔段 l_4 应在笔段 l_8 之前.

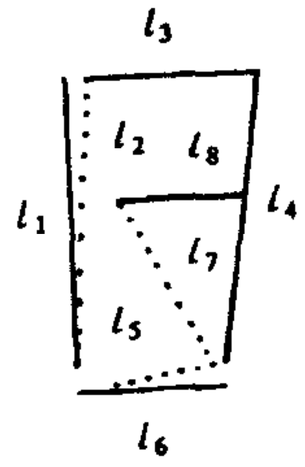

图 3　汉字“日”笔段分解

根据以上规则，得出构成图 3 中子结构“日”的笔段串为 $l_1l_2l_3l_4l_8l_7l_5l_6$. 应该指出的是，上述两条规则只能对部分笔顺化进行调整，子结构的匹配总的来说还是适合于固定笔顺的笔划串.

4.2　剩余部件的抽取

剩余部件是指待识汉字笔段串及候选字的参考模板笔段串中去除子结构笔段串后剩下的笔段串. 由于参考模板的子结构具有固定数目的笔划，因而很容易去除组成子结构的笔段而得到剩余部件，剩余部件的抽取主要针对待识汉字而言.

由于连笔的存在，当根据参考模板提供的笔划数抽取初始笔段串时，在字首子结构的末尾及字尾子结构的开始有可能包含一些剩余部件的笔段. 因而在待识汉字初始子结构笔段串与对应子结构的参考模板笔段串匹配完成后，子结构笔段串即剩余部件可根据以下规则动态确定：

1)将字首子结构的初始笔段串根据笔段书写的先后顺序重新排列，从笔段串的末端

开始往串首搜索，直到碰到能从参考模板子结构中找到对应笔段的笔段为止. 排列于该笔段之前(包括该笔段)的子笔段串为子结构笔段串，整字中排列于其后的笔段组成剩余部件.

2)将字尾子结构的初始笔段串根据笔段书写的先后顺序重新排列，从笔段串的开始往串尾搜索，直到碰到能从参考模板子结构中找到对应笔段的笔段为止. 排列于该笔段之后(包括该笔段)的笔段组成子结构笔段串，整字中排列于其前面的笔段组成剩余部件.

4.3 结构匹配

对每一个笔段，定义相应的绝对及相对属性. 绝对属性包括笔段相对与部件(子结构或剩余部件匹配时)或整字起点的坐标、笔段长度及运笔方向等，笔段的相对属性包括它相对于其它笔段坐标的相对顺序、各方向的笔段穿刺数等[4]. 由于子结构匹配的结果直接影响着整字的匹配，且子结构中各笔段的顺序相对稳定，因此，先对子结构笔段中进行排序，然后通过带有分裂-合并操作的动态规划法进行匹配[1]，对于子结构匹配后的剩余部件及整字，则将其笔段根据书写顺序排列成笔段串，通过准松驰法进行匹配[6].

5 实验结果

汉字库中的一级国标包含 3,755 个常用汉字，本文利用 35 套规整楷体样本对系统进行训练，建立标准模板，然后利用另外 3 套中等工整程度(有部分连笔，笔顺变化及书写倾斜)的一级国标样本进行测试，粗分类算法为每个待识汉字提供 10 个候选，为测试本文提出的子结构引导的结构匹配算法的性能，进行了以下测试：

1)为了检测活动模板对系统性能的影响，进行了两项实验. 首先，在对每类汉字的子结构属性进行人工训练(手工赋值)后，从具有相同子结构属性的汉字集中自动学习出一个对应子结构的标准模板，利用这些标准模板对待识汉字进行子结构检测，检测正确率仅为 86%；利用活动模板进行子结构检测，正确率达 98%.

2)利用本文方法，平均每个待识汉字需与 10 个侯选中的 5.1 个汉字匹配，而单纯的整子匹配法显然需于全部候选进行匹配，匹配运算量降低了约 50%.

3)利用整字匹配的联机系统的识别率为 93.9%，利用子结构引导匹配后，识别率基本不变. 这主要是因为利用子结构引导匹配算法虽然对汉字形变及笔顺变化的免疫力增强了，但子结构匹配本身的正确率并非 100%，产生累积误差，在加上前 10 候选的累积识别率就仅有 97%，识别率就难有很大提高了.

6 结论

本文在特征匹配与结构匹配综合的基础上，提出了一种子结构引导的结构匹配识别方法. 通过引入合适的字首子结构及字尾子结构，利用子结构检测的结果来引导汉字笔段串的匹配，大大减少了结构匹配运算，通过为子结构建立活动模板及设计子结构动态抽取算法，使得子结构匹配的准确度得到很大提高. 在联机手写汉字已取得了很大成就的今天，联机识别方法的突破有赖于对人脑智能模拟的水平，子结构的检测及各子结构间关系的描述与识别就是对人类识字过程的一种模拟，本文的方法对基于子结构检测的物体识别方法是很有启发意义的.

参 考 文 献

1 Tsay Y T, Tsai W H. Attributed string matching by split-and-merge for on-line Chinese character recognition. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1993, **15**(2): 180—185

2 刘迎建,戴汝为. 在线手写汉字识别的字形结构排序法. 自动化学报, 1988, **14**(3): 10—17

3 刘迎建,戴汝为. 联机手写汉字识别的理论与实践. 中文信息学报, 1989, **2**(4): 18—30

4 Xiao Xuhong. On-line handwritten Chinese character recognition by merging image features with Dynamic information. the 2nd China-Korea Joint Symposium on Oriental Language processing, 1996, 65—70

5 Kasturi, Raman *et al*. Document image analysis: an overview of techniques for graphics recognition. pre-proceedings of int. Workshop on Syntactic & Structural Pattern recognition, 1990, 192-230

6 Xiao Xuhong, Dai Ruwei. On-line Chinese signature verification by matching dynamic and structural features with a quasi-relaxation approach. in: Proc. 5th International Workshop on Frontiers in Handwriting Recognition, Colchee-tel England 1996

ON-LINE HANDWRITTEN CHINESE CHARACTER RECOGNITION GUIDED BY COMPONENTS WITH DYNAMIC TEMPLATES

XIAO XUHONG　DAI RUWEI

(Institute of Automation, The Chinese Academy of Sciences, Beijing　100080)

Abstract　Structural matching is a main approach for on-line Chinese character recognition. In order to reduce its great computational complexity and improve its performance, people have been seeking for a way to guide the whole matching by the result of partial matching. In this paper, the authors proposed 45 basic components from 3,755 categories of the daily-used Chinese characters to guide the stroke segment matching. Because they always locate at either the beginning or the end of the stroke segment string, these components are easy to extract and separate from other parts of the character. Besides, the reference templates of these components are dynamically extracted from the reference segment string and dependent on the current matched character so that a more accurate matching is carried out. Experiments show that the segment matching computation has reduced almost 50%. The approach is also enlightening for other similar object matching problem.

Key words　On-line bandwritten character recognition, structural matching, sub-structure, dynamic template.

肖旭红　1992 年毕业于中国科技大学无线电电子学系,1997 年于中国科学院自动化所获博士学位. 主要研究领域为:模式识别、人工智能、图象处理等.

戴汝为　主要从事人工智能、模式识别等研究,1980-1982 在美国普渡大学电机系作访问学者,现任中科院自动化所研究员,博士生导师,中科院院士,《模式识别与人工智能》主编,并任清华大学、汕头大学等近三十所高校的兼职或名誉教授.