

# 基于多 Agent 系统的脱机手写体汉字识别

马少平<sup>1,2,3</sup>, 金奕江<sup>1,2,3</sup>

(1. 清华大学 计算机科学与技术系, 北京 100084; 2. 清华大学 智能技术与系统国家重点实验室, 北京 100084; 3. 清华大学 清华信息科学与技术国家实验室(筹), 北京 100084)

**摘要:**由于脱机手写体汉字的多样性和随意性, 识别起来具有很大的难度, 依靠单一的特征很难实现高准确率的识别. 引入多 Agent 的概念, 将多种知识统一于多 Agent 系统之中, 给出了一个面向脱机手写体汉字识别的多 Agent 类市场模型, 提出了一种模糊综合方法和辩论协商规则, 实现了一个基于多 Agent 系统的脱机手写体汉字识别系统. 初步测试结果显示出系统的有效性.

**关键词:**汉字识别; 多 Agent 系统; 类市场模型; 模糊综合; 辩论协商规则

**中图分类号:**TP391.4 **文献标识码:**A **文章编号:**1673-4785(2009)05-0398-08

## Offline recognition of hand-written Chinese characters based on a multi-Agent system

MA Shao-ping, JIN Yi-jiang

(1. Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China; 2. State Key Lab of Intelligent Technology and Systems, Tsinghua University, Beijing 100084, China; 3. Tsinghua National Laboratory of Information Science and Technology, Tsinghua University, Beijing 100084, China)

**Abstract:** Due to the diversity and randomness of Chinese characters, it is difficult for offline hand-written Chinese character recognition to perform well when based solely on analysis of a single feature. In order to solve this problem, a multi-Agent based recognition method was proposed. It merges a variety of knowledge into a market-like model. A comprehensive approach using fuzzy rules to provide consultation and debate rules between Agents was also incorporated. With this proposed method, a multi-Agent offline hand-written Chinese character recognition system was constructed. Preliminary experimental results showed the effectiveness of this system.

**Keywords:** hand-written chinese character; multi-Agent system; market-like model; fuzzy synthetic; debate-negotiation rules

从 1966 年 IBM 公司首次发表关于汉字识别的研究文章以来, 汉字识别的研究已经有 40 多年的历史. 在这期间, 无论是在汉字特征的描述<sup>[1]</sup>、抽取方法, 还是在分类器的构造方法上, 均取得了很大进展, 尤其是在印刷体汉字识别和联机手写体汉字识别方面, 已开发出一批实用的系统<sup>[2]</sup>. 同时人们也清醒地认识到, 任何特征和分类器都有其局限性, 使用单一的特征、单一的分类方法, 很难使识别性能在现有的基础之上再上一个新的台阶. 为此, 人们又提出了综合使用不同的特征和分类器的系统集成方

法<sup>[3]</sup>, 以弥补单一特征、单一分类器的不足. 在印刷体汉字和手写体数字识别中, 已成功地开发出系统集成识别系统, 取得了良好的效果, 不仅降低了系统的误识率, 也使得系统的整体识别率达到了任何参与集成的单一系统所达不到的水平<sup>[4,5]</sup>.

在现有的系统集成方法中, 几乎无一例外地是多种识别器间的集成, 这种方法在现有条件下, 对脱机手写体汉字识别来说存在较大的困难, 以及不全面性和不合理性. 在多识别器集成系统中, 无论是采用简单的投票表决方法, 还是用各种不同的基于概率的决策模型, 均要求各分类器及所用特征间具有独立性, 这样才能使得决策结果具有公正性. 而目前较为有效的脱机手写体汉字识别方法中, 虽然各有特色, 但几乎全部是以方向线索特征<sup>[6]</sup>为基础的,

收稿日期: 2009-09-21.

基金项目: 国家自然科学基金创新研究群体科学基金资助项目 (60621062).

通信作者: 马少平. E-mail: msp@tsinghua.edu.cn.

不能满足多识别器合成的独立性要求。

手写汉字随意性较大,单靠特征很难正确地识别手写相近字,在印刷体中一些明显的差别,对于手写体来说也变地不那么可靠了。因此,单纯用多识别器集成来识别手写汉字,既不全面,也不尽合理,应使用除特征以外的更多的知识,进行综合判断。而在实际应用中,多是以文本为识别单位,这为综合利用多种知识提供了条件。后处理方法就是在这种情况下应运而生的<sup>[7]</sup>。

以往脱机手写体汉字识别模型的一个重要特点是其串行性,后处理引起的任何错误都将被保存下来,相对于识别来说,后处理具有权威性。而后处理的知识不可能是完备的,由于其权威地位,因此知识不足造成的错误得不到更正的机会。一种合理的想法是各种求解方法,无论是识别、后处理,还是其他方法,均处于一个平等的地位。不同的方法之间,通过协商、交流,达到总体上的和谐和最好的识别效果。多 Agent 系统正好具备这方面的特点<sup>[8-11]</sup>。本文引入 Agent 的概念,针对手写体汉字识别的特点,给出了一个基于多 Agent 系统的汉字识别模型,提出了一种模糊综合评判方法,以及各 Agent 间的协商辩论方法。

## 1 基于多 Agent 的汉字识别类市场模型

类市场模型是多 Agent 系统中的一个重要模型,受到学者们越来越多的重视。根据计算生态学的研究发现,在所有生态系统中,市场结构最为完善,具有最高的智能。根据手写体汉字识别的特点,借用市场模型的概念,给出基于多 Agent 系统的手写体汉字识别模型如图 1 所示。其中环境用于存放各种任务请求、处理所需的原始信息、处理的中间结果及其各 Agent 间的通讯信息等。管理者负责组织和规划任务求解以及环境中的信息管理。Agent 从环境中感知信息,利用自己的能力对其进行处理,并将处理结果通过管理者有组织地对环境进行更新。在该模型中,用户相当于消费者,向系统提出任务请求并接收系统的处理结果。管理者起中间商的作用,对欲求解的任务组织招标,规划中标者的任务求解,当出现矛盾冲突时,组织相关的 Agent 进行辩论和协商,以寻求一个可以接受的一致解。Agent 相当于生产者,时刻监视环境中的信息,一旦发现自己可以胜任的任务,则根据自己的能力进行投标,一旦中标后,立即进行求解。除了用户提供的任务请求外,每个 Agent 也可以根据自己的需求提出任务请求,以寻

求其他 Agent 的帮助。这样一来,Agent 在担当生产者的同时,有时也充当消费者的角色。

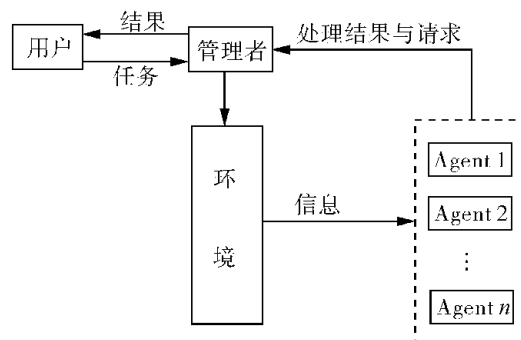


图1 汉字识别类市场模型

Fig.1 A market-like model for Chinese character recognition

系统的基本工作过程如下:

- 1) 用户向系统提交任务,任务中包括识别对象和要求等信息;
- 2) 管理者根据用户提交的任务,按用户的需求规划任务求解,形成招标信息公布于环境中;
- 3) Agent 感知到招标信息后,根据自己的求解能力及资源消耗,产生一种或几种投标方案,形成标书,送交管理者;
- 4) 管理者根据投标情况选择中标者,并规划任务求解,通过环境将中标信息分发给相应的 Agent;
- 5) Agent 在得到中标信息后,按中标合同组织问题的求解,将结果送交管理者,在结果中也可能包含新的任务请求;
- 6) 管理者对各个 Agent 的求解结果进行分析,若有新的任务请求,则组织招标,重复3~5),否则将结果进行组织后公布在环境中;
- 7) 在所有任务求解完成之后,仲裁 Agent 对识别结果进行仲裁,当存在矛盾冲突时,由管理者组织相关 Agent 进行辩论协商,直到达成一致解或最大可能解;
- 8) 最终由管理者将识别结果输出给用户。

## 2 模型的详细说明

### 2.1 任务

任务有2种类型,一种是由用户提交的任务,称为用户任务,它指定了系统的最终求解目标。另一种是由系统内部产生的任务,称为系统任务。管理者根据用户任务的不同目标,规划求解路径,路径中的每一个求解目标均产生至少一个系统任务。当 Agent 在求解过程中请求其他的 Agent 支援或需要与别的 Agent 对话时,生成系统任务。当发生矛盾冲突需要

进行辩论协商时,管理者也生成系统任务。

一个任务具有如下的格式:

( <任务名> <对象> <类型> <目标> [ <领域> ] )。

其中, <任务名> 是任务标识,由系统自动分配,任务名可以惟一地表示出一个任务; <对象> 给出该任务的操作对象,如一个切分任务,其对象可以是一个 TIF 或 BMP 文件,而一个识别任务,其对象则可以是已切分好的汉字图像点阵等; <类型> 指定了 <对象> 的类型,如 TIF\_FILE、BMP\_FILE 等; <目标> 给出该任务对给定对象最终求解到什么程度,如对于一个 TIF 文件对象,是只给出切分结果(因为用户可能并不要求识别)就行了呢,还是要对其进行识别,如果是识别是否进行后处理等等; <领域> 是一个可选参数,如果需要或者可能的话,它指定出 <对象> 所在的领域。领域既可以标识出识别对象是汉字、数字还是英文等信息,又可以给出待处理的对象属于社会科学范围,还是属于计算机科学范围等信息,供与领域有关的 Agent 使用。

## 2.2 环境

环境由一个公告牌和一个分层结构的黑板组成。用户提供的原始信息、招标投标信息、各 Agent 的处理结果及相互间的交互信息等均存放于环境之中。环境对于每个 Agent 是共享的。

公告牌是各种消息的集合。管理者与 Agent 之间、Agent 与 Agent 之间的各种通讯与交互均通过公告牌进行。

黑板是问题的解空间以层次结构方式组织起来的全局数据库,是所有公有信息的集合,Agent 使用的所有数据均存放于黑板之中。一个黑板被划分为以下 7 个层次:

1) 版面层。这是由扫描仪扫描汉字样张得到的最原始黑白二值图像,存储格式为 TIF 文件格式或 BMP 文件格式。最基本的可操作单位为图像的“点”。该层内容简称为版面。

2) 样本层。对版面进行分析后,经行切分、字切分后,得到单个汉字的点阵及其结构属性信息(如上下结构、左右结构、内外结构等)存放于该层。对汉字点阵进行噪声处理、光滑处理、规格化等预处理的结果也放于该层之中。汉字点阵基本的存储格式为二维矩阵,最基本的可操作单位为“点”。该层内容简称为样本。

3) 特征层。对汉字样本抽取出的识别特征存放于特征层,一个汉字的特征为一个  $N$  维向量,基本

的存储格式为数组。最基本的可操作单位为特征分量。一个汉字样本,可以对应多组特征。

4) 单字层。识别器基于特征的识别结果,连同其候选字、识别信度存放于该层。基本的存储格式为一结构,该结构含有 2 个关键域,一个域为按信度大小顺序存放的候选字,另一个域为候选字所对应的信度,2 个域均为数组。该层最基本可操作单位为字。

5) 词汇层。利用词汇知识对单字识别结果(包括候选)进行评判,评判结果存放于该层。基本的存储格式为含有 2 个关键域的结构,一个域记录候选字的构词情况,另一个域记录评判信度。该层的最基本可操作单位为词,包括单字词和多字词。

6) 短语层。对各候选可能形成的短语或句子,利用汉语的上下文知识进行评判,评判结果存放于该层。基本的存储格式是一个复杂的多级链表结构,实际上表达的是一个搜索图。该层的基本可操作单位为短语或句子。

7) 结果层。该层记录系统最终的识别结果,基本的存储格式同单字层一样为一结构,该结构含有 2 个关键域,一个域为按综合评判信度大小顺序存放的候选字,另一个域为候选字所对应的信度,2 个域均为数组。

## 2.3 管理者

管理者可以看作是一个特殊的 Agent,它具有多重身份。其一,管理者是一个中间商,它对用户或其他 Agent 提交的请求,规划求解路径,分解为若干个子任务,发布于公告牌上,组织招标。在接收到 Agent 的标书后,根据任务的具体要求,从求解精度、时间消耗和资源消耗等几个方面选择中标者。对于同一个求解目标,中标者可以是一个,也可以是多个。为发挥更多的 Agent 的作用,在时间允许的情况下,管理者尽可能多地选择中标者。其二,管理者实现对环境的管理。所有的任务请求,均通过管理者张贴于公告牌上,所有处理结果,也要经过管理者组织之后放置于环境之中。其三,管理者是一个调节人,当发生矛盾冲突时,管理者负责组织各相关 Agent 间的辩论与协调,听取各辩论者的意见,使得在各 Agent 间最终达成一个一致的意见或可能性最大的结果。

## 2.4 Agent

Agent 由感知器、发送器、任务分配器、知识库、方法集和局部黑板 6 部分组成,其一般结构如图 2 所示。其中感知器用于感知环境中的信息,它时刻监视着环境的变化,随时捕捉与自己相关的信息。任务

分配器根据感知到的信息,分发给适合的方法.方法集是该 Agent 能力的体现,与知识库相配合,实现对问题的求解.方法集至少由 3 部分内容组成:投标方法、问题求解方法和辩论协商方法.方法集和知识库构成了 Agent 的大脑,是 Agent 最重要的组成部分.

局部黑板是 Agent 的私有数据库,用于存储求解问题所需的各种数据、中间结果及最终结果等.发送器将求解结果或任务请求发送给系统的管理者,以实现与环境或其他 Agent 的交互.

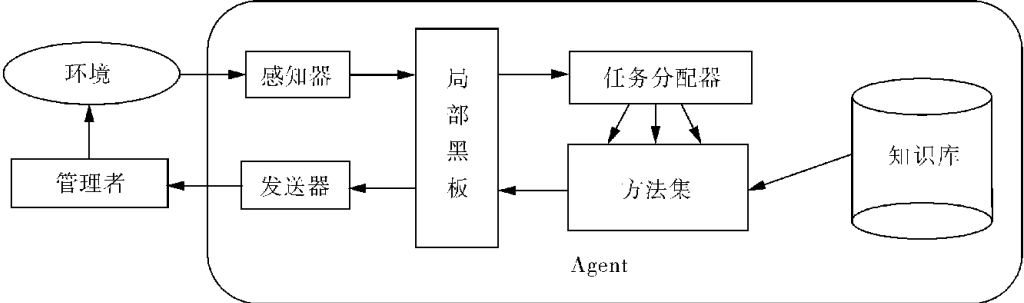


图 2 Agent 的一般结构  
Fig.2 General structure for recognition Agent

Agent 根据其功能的不同,可以分为以下几类:

- 1)扫描 Agent:启动扫描仪,获得待识别文字的图像信息.
- 2)切分 Agent:对版面进行分析,将版面中的每个汉字从图像中分离出来,得到待识字样本;必要时,切分 Agent 也可以给出样本的结构信息,如左右结构、上下结构等.
- 3)预处理 Agent:消除样本中存在的噪声,对汉字笔画边缘进行平滑处理,然后再对汉字第本进行非线性整形变换及大小归一化处理.
- 4)特征抽取 Agent:从归一化后的汉字第本中抽取识别用特征.
- 5)识别 Agent:对于不同的特征,采用不同的方法对待识字样本进行分类,得到候选字及其识别参数.
- 6)识别评价 Agent:应用单字识别系统的误差模型及识别参数对候选字进行评价,得到候选字的识别信度.
- 7)词汇处理 Agent:利用词汇知识对前后相关联的候选字进行构词分析,提出假设,并给出信度评价.
- 8)后处理 Agent:对各候选可能形成的短语或句子提出假设,利用汉语语言模型进行分析,给出信度评价.
- 9)仲裁 Agent:对不同的 Agent 给出的结果,用某种评判方法进行综合评判,一致的部分确定下来,产生矛盾的部分,送交管理者组织辩论.
- 10)自动校对 Agent:对识别结果中与语言模型不相符的部分提出警告,提示给用户.
- 11)人工校对 Agent:提供一种便于用户校对识

别结果的手段.

各类 Agent 与环境信息层的关系如图 3 所示.

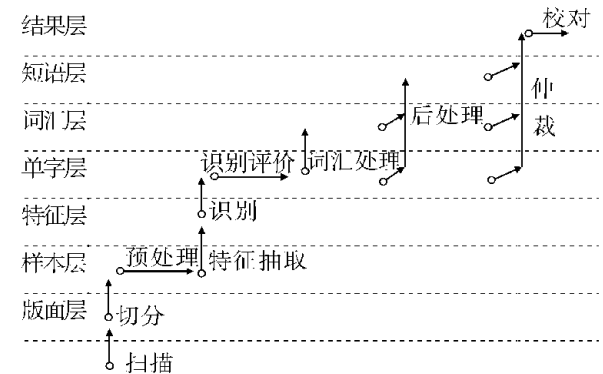


图 3 知识源与信息层的关系  
Fig.3 Relationship between knowledge sources and information layers

2.5 仲裁

对于脱机手写体汉字识别来说,各 Agent 的处理结果很难做到完全一致,当出现任何不一致时,系统就进行辩论协商,系统开销太大.一种可行的办法就是对结果进行模糊综合评判,当评判结果达到一定的可信度时,就认为该结果是一致的,否则被认为是有矛盾冲突的.只有在模糊评判意义下发生冲突时才进行辩论协商.

2.6 协商与辩论

协商是多 Agent 系统中关键的组成部分<sup>[12]</sup>.若若干个 Agent 简单地堆放在一起,永远是几个独立的个体,只有相互协调合作,才能使其综合能力具有质的变化.辩论是协商的一种方式,通过辩论,使得各 Agent 间取得一致的意见,也就是说,得到一个对于

待识别样本可能性最大的识别结果。

辩论是一个说理的过程,每个参与辩论的 Agent,从自己的立场出发,重新审视所讨论的问题,提出自己的理由和根据。通过协商,或者坚持自己的原有观点,努力去说服其他的 Agent 同意自己的意见;或者被其他的 Agent 说服,改变自己的立场,支持说服者的意见。下面给出一些辩论规则。

规则1 如果  $C_1$  是特定结构汉字,而  $C_2$  是非特定结构汉字,当当前待识别的汉字  $O$  与  $C_1$  具有相同的特定结构时;则有理由相信  $O$  为  $C_1$ ,而非  $C_2$ 。

其中特定结构汉字指的是具有左右结构,或者上下结构,或者内外结构的汉字。

$C_1$ 、 $C_2$  的结构信息存放于 Agent 的知识库中,而  $O$  的结构信息在切分时获得,或者通过求解笔画的连通域获得。

注意该规则只规定了当  $O$  与  $C_1$  具有相同的特定结构时才确信  $O$  为  $C_1$ ,而  $O$  为非特定结构时,并不能确信  $O$  为  $C_2$ 。这是因为在手写体汉字中,习惯性的连笔往往会破坏汉字的结构特征。

规则2 如果  $C_1$  与  $C_2$  的复杂度差大于给定值,且  $O$  与  $C_1$  的复杂度差小于给定值;则有理由相信  $O$  为  $C_1$ 。

其中汉字的复杂度可以用规格化后汉字点阵的黑白点之比来度量,也可以用汉字的纵向层次数或横向层次数度量。

规则3 如果  $O$  为汉语句子里的一员,且  $C_1$ 、 $C_2$  为一对客观相似字,则识别类 Agent 对此不参与辩论。

将相似字定义为客观相似字和主观相似字<sup>[13]</sup>,客观相似字指的是那些拓扑结构相似的汉字,而主观相似字指的是那些拓扑结构并不太相似,但是特征比较相似的汉字,这是由于特征抽取的不连续性造成的。

客观相似字虽然它们的拓扑结构非常相似,但字意一般有很大的差别,如“土”和“土”、“末”和“未”等。与其通过识别 Agent 找出它们在字形上的差别,不如通过后处理等手段进行选择更为可靠。

规则4 如果  $P$  是使  $C_1$ 、 $C_2$  产生最大差异的部分,且在相同的部分  $O$  与  $C_1$  的差异小于  $O$  与  $C_2$  的差异;则有理由相信  $O$  为  $C_1$ 。

其中,  $P$  为汉字的左半部、右半部、上半部、下半部、中心部或外围部之一。差异指的是当只使用包含

在  $P$  部的特征进行识别时的识别器评价指标。

规则5 如果  $P_1$  是相对于 Agent1 的  $C_1$ 、 $C_2$  间的最大差异部分,  $P_2$  是相对于 Agent2 的  $C_1$ 、 $C_2$  间的最大差异部分,  $P_1 \neq P_2$ , 且当用  $P_1$  替换  $P_2$  时 Agent2 支持 Agent1 的结论;则有理由相信 Agent1 的结论是正确的。

该规则反映了 Agent 在辩论中的退让。

规则6 如果已经确认  $O$  为专有名词的一员,且选择  $C_1$  后,专有名词词典中含有该名词;则有理由相信  $O$  为  $C_1$ 。

专有名词指人名、地名、公司名等,在一个句子中,具有明确上下文特征的专有名词可以通过判别法则判定。

规则7 如果已经确认  $O$  为专有名词的一员,且选择  $C_1$  或  $C_2$  后,在专有名词词典中均不含有该名词;则识别类 Agent 的结果更为可靠。

规则8 如果识别 Agent 对  $C_1$  和  $C_2$  的可信度之差小于给定值,而且  $C_1$  的组句能力与  $C_2$  的组句能力之差大于给定值;则有理由相信  $O$  为  $C_1$ 。

其中  $C$  的组句能力定义为:当  $O$  固定为  $C$  时,经处理后  $O$  所在句子的概率。这也是一条退让规则,反映了当识别 Agent 没有较大的把握区分出  $C_1$  和  $C_2$  时,把决定权交由后处理 Agent。

### 3 模糊综合评判

设  $x$  为待识汉字,  $C = \{c_1, c_2, \dots, c_n\}$  为其候选集,  $A = \{a_1, a_2, \dots, a_m\}$  为 Agent 集,每一个  $a_j$  从自己的立场出发,对候选  $c_i$  是否为  $x$  的识别结果有一个评价,经适当的转换后,该评价可以看作是从  $A$  到  $F(C)$  的模糊映射,即

$$\tilde{f}: A \rightarrow F(C),$$

$$a_j \mapsto \tilde{f}(a_j) \triangleq (r_{j1}, r_{j2}, \dots, r_{jn}) \in F(C).$$

式中:  $F(C)$  表示定义域为  $C$  的模糊集合的全体。

由于各 Agent 所采用的知识不同,其对  $c_i$  评判的精确程度和重要程度也不同,因此对不同的 Agent 给出的评判,要分别对待,有一定的权重分配。权重分配可以看作是  $A$  上的模糊集,记为

$$W = (w_1, w_2, \dots, w_m) \in F(A).$$

式中:  $w_j$  表示第  $j$  个 Agent 的权重,它们满足归一化条件:

$$\sum_{j=1}^m w_j = 1.$$

对各候选综合评判的结果,可视为  $C$  上的模糊集,记为

$$B = (b_1, b_2, \dots, b_n) \in F(C).$$

式中:  $b_i$  反映了综合评判后,待识字  $x$  被识别为  $c_i$  的可能程度.

由模糊映射与模糊关系的关系及模糊关系与模糊变换的关系,模糊映射  $\tilde{f}$  可以惟一地诱导出模糊关系:

$$R \triangleq R_{\tilde{f}} \triangleq \begin{bmatrix} \tilde{f}(u_1) \\ \tilde{f}(u_2) \\ \vdots \\ \tilde{f}(u_m) \end{bmatrix} \triangleq \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix} \in M_{m \times n}.$$

式中:  $M_{m \times n}$  为  $m \times n$  矩阵的全体.

由  $R$  又可以惟一地诱导出一个模糊变换:

$$\tilde{T}_R: F(A) \rightarrow F(C),$$

$$W \mapsto \tilde{T}_R(W) \triangleq W \circ R.$$

这样,由三元组  $(A, C, R)$  构成了一个模糊综合评判模型. 当给定一个权重分配  $W = (w_1, w_2, \dots, w_m) \in F(A)$  后,则输出一个模糊综合评判  $B = W \circ R = (b_1, b_2, \dots, b_n) \in F(C)$ , 于是有

$$(b_1, b_2, \dots, b_n) = (w_1, w_2, \dots, w_m) \circ$$

$$\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ r_{21} & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ r_{m1} & r_{m2} & \cdots & r_{mn} \end{bmatrix}.$$

如果  $b_0 = \max\{b_1, b_2, \dots, b_n\}$ , 则评判  $c_0$  为  $x$  的识别结果. 式中算子“ $\circ$ ”可取 Zadeh 算子  $(\wedge, \vee)$ 、概率算子  $(\cdot, +)$  或其他模糊算子. 采用不同的算子,所强调的内容不同,评判的结果也有所不同.

当采用 Zadeh 算子时,

$$b_j = \bigvee_{i=1}^m (a_i \wedge r_{ij}), j = 1, 2, \dots, n.$$

式中:

$$x \wedge y = \min(x, y),$$

$$x \vee y = \max(x, y).$$

当采用概率算子时,

$$b_j = w_1 \cdot r_{1j} + w_2 \cdot r_{2j} + \cdots + w_m \cdot r_{mj}, \\ j = 1, 2, \dots, n.$$

式中:

$$x \cdot y = x \times y,$$

$$x + y = x + y - x \times y.$$

Zadeh 算子强调单个 Agent 的作用,当有一个 Agent 对某个候选表现出很强的可信度时,就将该候选作为评判结果. 而概率算子则考虑了更综合的情况,具有一定的加权平均的意义.

由于各 Agent 对候选的评判是基于统计知识进行的,同时又存在着“数据稀疏”问题,所以这种评判带有很大的先验性和不准确性;因此在评判时采用具有加权平均意义的概率算子比用主元素决定的 Zadeh 算子要好一些,经实际测试也证明了这一点. 而且这与多 Agent 系统发挥各 Agent 的综合优势也是相吻合的.

## 4 系统测试

以上文所述模型为基础,实现了一个基于多 Agent 的手写体文本汉字识别系统,对系统的性能进行了测试. 为了更好地体现多 Agent 的作用,特意选取了一个早期的识别率相对比较低的脱机手写体汉字单字识别系统,以体现多种知识融合的作用.

测试条件如下:

1) 训练样本: 中科院自动化所样本库, 共含国标一级汉字 50 套;

2) 汉语语料库:《人民日报》语料库, 共含汉字 1 500 万个;

3) 识别库大小: 1.9 MB;

4) 后处理库大小: 4.5 MB;

5) 文章内容: 从《人民日报》、《新清华》上选取的 5 篇文章, 每篇文章的字数从 770 字到 2 855 字不等;

6) 书写对象: 随机选取的 6 位书写者;

7) 书写要求: 按自己习惯工整手写.

采用以下指标测试系统的性能:

$$\text{单字识别率} = \frac{\text{单字识别正确字数}}{\text{总字数}} \times 100\%,$$

$$\text{综合识别率} = \frac{\text{综合正确字数}}{\text{总字数}} \times 100\%,$$

$$\text{综合校正率} =$$

$$\frac{\text{综合正确字数} - \text{单字识别正确字数}}{\text{总字数} - \text{单字识别正确字数}} \times 100\%.$$

式中:“单字识别正确字数”指的是只采用识别 Agent 时,系统能正确识别的汉字数;“综合正确字数”指的是采用多 Agent 之后,系统总的正确识别的汉字数.

综合校正率是反映多 Agent 汉字识别系统性能的一个重要指标,它反映了采用多 Agent 之后系统误识率的下降程度.

测试结果如表1所示,从表中可以看出,系统的平均单字识别率为92.2%,而综合识别率平均达到了98.1%,提高近6%,综合校正率平均为74.2%。

表1 文本识别实验结果

Table 1 Experimental results on character recognition

书写者	字数	单字 识别率/%	综合 识别率/%	综合 校正率/%
1	770	91.0	98.0	77.8
2	1 533	91.7	98.5	81.9
3	2 855	90.3	97.0	69.1
4	1 511	89.9	97.8	78.2
5	1 511	96.1	98.6	64.1
6	1 522	94.6	98.6	74.1
平均		92.2	98.1	74.2

部分单字识别结果与系统的综合处理结果对比例句如下。其中:“输入文字”是手写汉字的内容;“识别结果”是单字汉字识别系统给出的识别结果;“综合结果”是系统综合以后得到的结果。有下划线的汉字为误识字。

1)输入文字:两国领导人为两国关系的现状感到满意。

识别结果:两目领导人讨西国类系约现状礁到满意。

综合结果:两国领导人为两国关系的现状熬到满意。

2)输入文字:在会谈中。

识别结果:礼会误干。

综合结果:在会谈中。

3)输入文字:他表示愿为保持和发展两国的友好合作关系作出自己的贡献。

识别结果:他先示叁杠探拱和发在丙国的辰好言作关系作出自己封贡献。

综合结果:他表示愿为保持和发展两国的展的合作关系作出自己时贡献。

4)输入文字:是建设有中国特色社会主义伟大事业的主力军。

识别结果:是建设有中国特邑社余立叉伟大事业的主力军。

综合结果:是建设有中国特色社会主义伟大事业的主力军。

5)输入文字:社会稳定了我们才能集中精力做好各项工作。

识别结果:社会稳定了我们才能菜中精才做好务项工作。

综合结果:社会稳定了我们才能集中精力做好合项工作。

6)输入文字:我自然是不赞成的。

识别结果:战自然足不赞或的。

综合结果:我自然是不赞成的。

7)输入文字:亿万农民解放思想更新观念。

识别结果:亿万农民解放怠想更新观念。

综合结果:亿万农民解放思想更新观念。

8)输入文字:更好地发挥工人阶级主力军的作用。

识别结果:更好地发挥工人阵级主力军钧诈同。

综合结果:更好地发挥工人阶级主力军的作用。

9)输入文字:要广泛宣传具有鲜明时代特征的艰苦创业先进典型。

识别结果:要广沧室传具有鲜明时伐特征的艰苦创业先迭典型。

综合结果:要广泛宣传具有鲜明时代特征的艰苦创业先进典型。

10)输入文字:九十年代实现翻两番的第一步战略目标。

识别结果:丸十年代卖观翻两番的第一步战暗目标。

综合结果:九十年代实现翻两番的第一步战略目标。

11)输入文字:是保证中心组学习活动的重要条件。

识别结果:是俘证中心狙学习玲劫钩重要条件。

综合结果:是保证中心组学习活动的重要案件。

从以上例句可以看出,系统具有非常强的综合校正能力,大多数识别错误经多 Agent 综合处理后,得以校正。有时连续出现的误字,也能被改正,如在例句3、4、8、11中,出现了连续3个以上的误字同时被更正的情况。

系统的误识主要是候选不足造成的。如例句1中“感到满意”的“感”字,例句3中“自己的贡献”中的“的”字等均未进入候选。同样在例句3中“友好合作”部分,由于“友”字未入候选,从而株连到“好”字,正字也变误字了。这种现象称为“误校”。误校的另一个原因是后处理的能力不足造成的。如在例句11中,“重要条件”是正确的,被误校为“重要案件”。其原因是,按照 Markov 模型,“重要案件”的连接概率大于“重要条件”的连接概率,而“条”与“案”字又比较相近,识别 Agent 容易做出退让。

## 5 结束语

本文引入 Agent 的概念,提出了一个基于多 Agent 系统的手写体汉字识别类市场模型。该模型改变了以往的“切分——识别——后处理”的串行处

理机制,各 Agent 处于同等的处理地位,每个 Agent 从各自的立场出发,提出自己的处理意见,通过模糊综合评判,给出一个模糊评判意义下的最优结果.当综合评判不能确定结果时,采用辩论的方法,协调各 Agent 的意见,通过 Agent 间的协商,得到一个大多数 Agent 都能接受的识别结果,任何一个 Agent 都不处于绝对权威的地位,充分发挥了每个 Agent 自身的决策能力,从而极大地减少了串行机制中,每一个局部的处理错误都将导致最终错误的不合理现象的发生.该模型通过招标投标的形式求解问题,具有一定程度上的开放性.组成系统的 Agent 的数量及功能均可以动态决定,系统随时可以加入新的 Agent 参与任务求解.从原理上讲,该模型具有分布处理能力,各 Agent 既可以集中在一台处理器上,也可以分布在不同的处理器上,为进一步的网络计算打下了基础.

## 参考文献:

- [1] 陈 静,穆志纯,孙筱倩. 计算机模拟汉字字形认知过程的研究[J]. 智能系统学报, 2008, 3(3): 216-221.  
CHEN Jing, MU Zhichun, SUN Xiaoqian. Computer simulation of the cognition of Chinese characters [J]. CAAI Transactions on Intelligent Systems, 2008, 3(3): 216-221.
- [2] 张忻中. 汉字识别技术[M]. 北京:清华大学出版社, 1992: 31-41, 125-160.
- [3] DU Qingdong, LIU Jie. A new neural fusion recognition method with multi-Agent[C]//Proceedings of the Third International Conference on International Information Hiding and Multimedia Signal Processing (IIH-MSP 2007). Washington DC: IEEE Computer Society, 2007: 127-130.
- [4] ZHU Xiaoyan. Multiple neural networks model and its application in pattern recognition[C]//IEEE International Conference on Neural Information Processing. Beijing, China, 1995: 966-969.
- [5] 张永慧,刘昌平,罗 公,等. 技术综合集成在模式识别中的应用[J]. 计算机学报, 1995, 18(19): 678-685.  
ZHANG Yonghui, LIU Changping, LUO Gong, et al. Integration comprehensive techniques in pattern recognition[J]. Chinese Journal of Computers, 1995, 18(19): 678-685.
- [6] 马少平,夏 莹,朱小燕. 基于模糊方向线索特征的手写体汉字识别[J]. 清华大学学报:自然科学版, 1997, 37(3): 42-45.  
MA Shaoping, XIA Ying, ZHU Xiaoyan. Handwritten Chinese characters recognizing based on fuzzy directional line element feature[J]. Journal of Tsinghua University: Sci & Tech, 1997, 37(3): 42-45.
- [7] 夏 莹,马少平,常新功,等. 基于统计的汉字识别文本的自动后处理方法[J]. 模式识别与人工智能, 1996, 9(2): 172-178.  
XIA Ying, MA Shaoping, CHANG Xingong, et al. The method of automatic post-processing based statistical probabilities for Chinese recognition text[J]. Pattern Recognition and Artificial Intelligence, 1996, 9(2): 172-178.
- [8] RUSSELL S, NORVING P. 人工智能——一种现代方法[M]. 姜 哲,金奕江,张 敏,等,译. 2版. 北京:人民邮电出版社, 2004: 26-42.
- [9] LIU Jiming. 多智能体原理与技术[M]. 靳小龙,张世武,LIU Jiming,译. 北京:清华大学出版社, 2003: 1-17, 43-64.
- [10] HUYNH T D, JENNINGS N R, SHOABOLT N R. An integrated trust and reputation model for open multi-Agent systems[J]. Autonomous Agents and Multi-Agent Systems, 2006, 13(2): 119-154.
- [11] PANAIT L, LUKE S. Cooperative multi-Agent learning: the state of the art[J]. Autonomous Agents and Multi-Agent Systems, 2005, 11(3): 387-434.
- [12] 王立春,陈世福. 多 Agent 多问题协商模型[J]. 软件学报, 2002, 13(8): 1637-1643.  
WANG Lichun, CHEN Shifu. A multi-Agent multi-issue negotiation model[J]. Journal of Software, 2002, 13(8): 1637-1643.
- [13] 张德喜,马少平,朱绍文,等. 基于统计与神经元方法相结合的手写体相似字识别[J]. 中文信息学报, 1999, 13(3): 33-39.  
ZHANG Dexi, MA Shaoping, ZHU Shaowen, et al. Handwritten similar Chinese characters recognition based on combining statistics with neural networks method[J]. Journal of Chinese Information Processing, 1999, 13(3): 33-39.

## 作者简介:



马少平,男,1961年生,教授,博士生导师,主要研究方向为智能信息处理、信息检索、汉字识别与后处理以及中文古籍数字化.承担过多项国家自然科学基金、“863”项目、“973”项目及国际合作项目,在脱机手写体汉字识别和后处理方面达到了国际先进水平,“脱机手写体汉字与数字识别系统”1998年1月获得国家教委科技进步二等奖.发表学术论文70余篇,出版专著2部.



金奕江,男,1970年生,工程师,主要研究方向为汉字识别、信息检索与处理.发表学术论文10余篇.