

基于互动计算的汉字楷书识别¹⁾

周昌乐

(杭州大学计算机系 310028)

马希文

(北京大学计算所 100871)

摘 要

本文从互动计算的角度出发,对汉字楷书识别问题作了专门的研究,给出了一个具体算法及其实现。由于强调了识别的整体性、关联性、相对性和同时性,有效地解决了汉字形似字的识别问题,识别率达80%以上。

关键词: 视觉分类,汉字识别,整体性,互动计算。

一、基本思想

汉字楷书识别是属于视觉分类问题。很明显,对于一个汉字整体的识别,必须通过对这一整体的各组成要素进行分析把握之后才能完成;然而,如果在分析时,没有一个整体的概念作指导,对这个整体汉字的识别就一步也不能深入下去。比如在手书“地”中,我们之所以将左下方的笔画认作提而不是横,把该字认作左右结构而不是其它结构,都是因其在该字中所处的地位、关系所决定的。更一般地,由于处理数据的策略往往只有在处理数据的同时才能确定,而且其确定又是一个不断回环的整体演化过程,所以要解决象汉字楷书识别这样的问题,就必须采用互动计算方法:不仅利用知识,更重要的是利用当前整合结果指导对当前数据的进一步整合,并重复这一过程直至得到最理想的整合结果。

二、算法的实现

如果用树结构表示识别过程在某时刻反映所给定汉字诸要素及其关系的当前整合结果,那么按互动计算的思想,识别给定汉字的具体过程便是从开始的线条要素,通过互动计算的逐步构造,最终寻找出符合汉字组字知识的最佳树结构并使其与预置的样本吻合。具体算法简述如下:

本文于1990年12月3日收到。

1) 本课题为北京大学视觉与听觉国家开放实验室赞助项目。

- 1) 输入汉字图形 v 并置当前指导值 u 为初值;
- 2) 根据 u 分解 v 为 n 条线条;
- 3) 以 n 条线条为叶节点形成识别树结构 P , 并置当前节点层计数 $h = 0$;
- 4) 置根节点指导值 $= u$;
- 5) 计算 P 及代价效果值 $cost$;
- 6) 根据组字规则和图式选择 P 当前节点层所有可能的构字组合为 l 种;
- 7) 如果 $l < 0$, 则(处理线条)
 - (a) 如果 P 叶节点内容均已完全确认, 则 Print (识认困难)转 16);
 - (b) 修正 u 转 2);
- 8) 如果 $l = 0$, 则(处理冲突)
 - (a) 修正 u ,
 - (b) 如果 P 根节点指导值 $= u$, 则 Print (识认困难)转 16), 否则转 4);
- 9) 对所有组合 $k(k = 1, 2, \dots, l)$ 同时进行
 - (a) 根据组合 k 及 P 构建新的树结构 P_k ;
 - (b) 计算 P_k 及代价效果值 $cost_k$;
- 10) 求出最佳树结构集 $K = \{k | \min(cost_k < cost)\}$;
- 11) 如果 $|K| \neq 1$, 则(处理竞争)
 - (a) 修正 u ;
 - (b) 如果 P 叶节点内容均已完全确认, 则 Print (识认困难)转 16), 否则转 2);
- 12) 如果 $l > 1$, 则(处理局部极小, k_0 为其下标)
 - (a) 修正 u ;
 - (b) 如果 P_{k_0} 与 P 一致且 P 根节点指导值 $= u$, 则 Print (识认困难)转 16);
 - (c) $P = P_{k_0}$, $h = h + 1$, 转 4);
- 13) $P = P_{k_0}$, $h = h + 1$;
- 14) 如果 $n_{k_0} > 1$, 转 5)(深化识别过程);
- 15) 如果结果样本不唯一, 则(处理微小差别)
 - (a) 针对具体每个样本计算 P 及代价效果值;
 - (b) 如果诸代价效果最小者不唯一, 且 P 叶节点内容均已完全确认, 则 Print (识认困难)转 16), 否则修正 u 转 2); 否则取代价效果值最小者为 P ;
- 16) 输出 P 根节点理解值, 结束.

其中 P 树结构节点用四元组〈理解值、要素位置、指导值、子要素个数〉表示, 除了用以反映组成汉字诸要素的特征属性(相对长度、相对走向、相对位置、要素代码以及指出连体、笔画、叉点、拐点的拓扑指数)外, 也为 P 的计算提供了依据, 并体现互动思想; 代价效果值的计算公式为:

$$\text{相容样本个数} = \max_{\text{相容样本}} \left\{ \frac{\text{相容样本与 } P \text{ 匹配一致的节点数}}{P \text{ 的节点总数}} \right\}.$$

例如, 对于楷书“地”, 按照算法, 可用图 1 给出其树结构的演变过程. 首先是得出六画字, 然后经连体唯一的组合方式进一步得出连体为 2、叉点为 3 的相容样本{存, 在, 戏,

托,地},最后经过二种可能组合的选择得到最终树结构并确认为“地”。当然,如果线条有

缺疑或遇到冲突、竞争、局部极小和微小差异时,算法还会通过修正指导值重新有针对性地去分割线条或重构组合,这里不再赘述。

结果说明

算法是在北京大学视觉、听觉国家实验室实现的。共计源程序1400行,样本字库统计收入四千常用汉字。调试结果令人满意,就笔者手书测试(随意书写的三千汉字)的统计及博士论文鉴定来看,识别率已达80%以上并且对书写没有限制,基本上实现了本文的设想。当然,存在的问题也不少,特别是对于区分不符合楷书书写规范的交错连笔、同向连笔现象往往出现严重失误,这是有待于深入研究的问题。

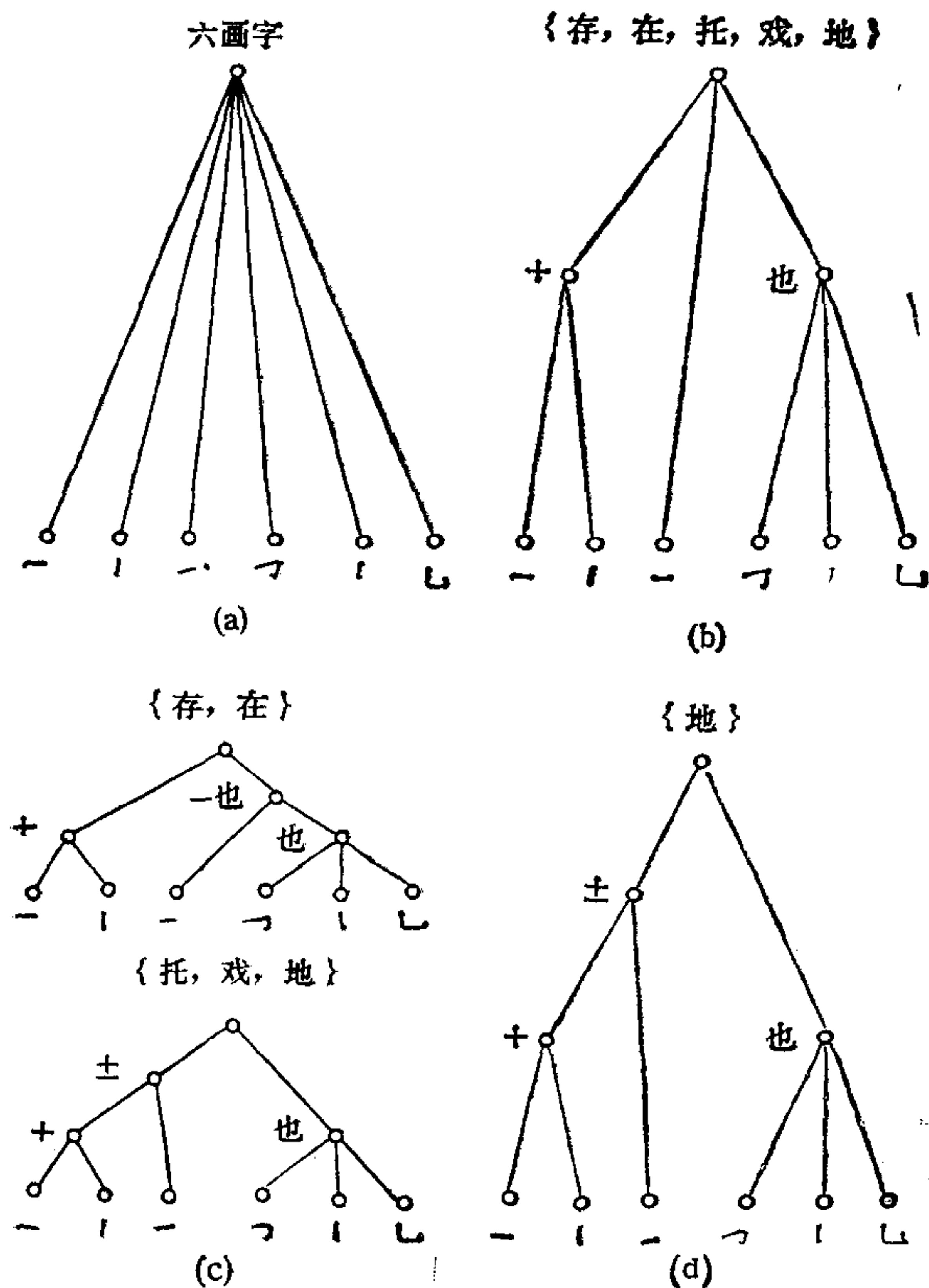


图1 “地”树结构的演变

RECOGNIZING REGULAR-SCRIPT CHINESE CHARACTERS BY AN INTERACTIVE COMPUTATION

ZHOU CHANGLE

(Dept. of Computer, Hangzhou University, 310028)

MA XIWEN

(Institute of Computing Science, Peking University, 100871)

ABSTRACT

From the view of interactive computation, the problems of the recognition of regular-script Chinese characters have been studied, and a computer recognition system has been constructed. Since attention has been paid to the entirety, relativity, simultaneity, and interrelationship in the recognition, the problem of how to distinguish similar characters can be effectively solved.

Key words: Classification of vision; recognition of Chinese characters; entirety; interaction.