

笔画节点在手写体汉字识别中的作用*

朱一鸣 赵阳 唐宁 周吉帆 沈模卫

(浙江大学心理与行为科学系, 杭州 310058)

摘要 产生式理论认为, 视觉图形的识别是对其产生过程的逆推理。汉字是笔画按正字法规则交错连接构成的象形文字, 手写体汉字识别可以认为是对汉字产生过程的反向推理。基于典型的产生式模型——贝叶斯规划学习模型, 汉字的产生式识别过程从识别字的笔画开始, 先基于线段交点提取出节点, 再枚举能产生该节点的所有笔画组合方式, 从而获得汉字的产生方式。据此预测, 节点数量和节点复杂度是手写汉字识别过程的重要影响因素。本研究通过三个实验考察了节点在汉字识别中的作用。结果显示, 含有较多节点的汉字具有更好的识别绩效(节点数量效应), 掩盖由较多笔画构成的高复杂度节点会对汉字识别产生更大干扰(节点复杂度效应)。本研究增进了对汉字识别早期过程的认识, 为字形识别的产生式反向推理过程提供了证据。

关键词 手写汉字识别, 节点, 笔画, 产生式模型

分类号 B842

1 前言

汉字是利用二维空间表达信息的象形文字, 由笔画交错连接构成, 其正字法规则(Orthographic regularities)较为复杂(陈天泉, 1983)。虽然当前印刷体汉字识别技术已经成熟, 手写体汉字因其笔画多变、风格各异, 给机器识别带来巨大挑战(任晓倩等, 2018; Krizhevsky et al., 2012)。然而, 汉字使用者往往具备熟练识别手写汉字的能力。因此理解并借鉴人的手写汉字识别机制, 探明笔画、部件等各层次单元的特征及计算机制, 有着重要的理论意义和应用价值(周吉帆等, 2016; Li et al., 2020; Zhang et al., 2020)。

植根于建构主义的视觉合成分析过程理论(Analysis-by-synthesis process)认为, 所有图像都是由某些因果过程在时间和空间上的执行而产生的, 动态生成的结果是一幅静态图像。面对这一输出结果, 人的视觉系统会自发地以概率计算的方式解释图像是如何产生的(Grenander, 1976; Yuille & Kersten, 2006)。例如, 人们将苹果公司的标志看作是被“咬

了一口”的残缺苹果。面对汉字, 人的视觉系统可能执行了类似的过程: 汉字是笔画和部件按照自上而下、自左向右等正字法规则书写生成的产生式文字, 汉字识别可能是根据整字这一静态输出结果, 逆向推理此前的动态书写过程, 猜测输入图像最有可能由哪个原型字书写而成。该产生式的概率推断思想得到行为学和神经科学证据的支持(Gershman et al., 2012; Kok et al., 2013; Laeng et al., 2014)。可见, 如果汉字识别是一个产生式过程, 那么首先需要理解汉字是如何产生的。

汉字是由内含层次关系的各个单元在正字法的引导下生成的。以往的模型大多将汉字分为笔画、部件、整字三层次, 例如经成分模型(Huang & Wang, 1992)、合体汉字识别的相互作用模型(沈模卫, 朱祖祥, 1997)、多层次交互模型(Taft & Zhu, 1997)和格式塔认知模型(陈传锋, 黄希庭, 2004)。尽管各模型有所差别, 但它们隐含了一致的观点, 即笔画是汉字结构和识别的最小单元。如果汉字识别的过程相当于汉字产生的反向推理, 那么获得笔画表征应当是汉字识别的前提。

收稿日期: 2022-01-14

* 国家自然科学基金面上项目(32071044、31871096)和中央高校基本科研业务费专项资金资助(2021FZZX001-06)。

通信作者: 周吉帆, E-mail: jifanzhou@zju.edu.cn; 沈模卫, E-mail: mwshen@zju.edu.cn

支持笔画是汉字识别最小单元的实验证据主要来自汉字的笔画数效应、笔顺效应和笔画独立组块效应。笔画数效应是指笔画数多的汉字加工更困难。研究发现,识别或辨认笔画数较多的合体字需要花费更多时间(张武田,冯玲,1992),同时正确率和辨别力更低(郑昭明,高尚仁,1982);判断一组笔画是否构成汉字时,少笔画组构字的错误率低于多笔画组(罗艳琳等,2008)。上述研究提示,整字包含的笔画数量越多,需要加工的特征量也越多,支持笔画是汉字字形加工的最小单元。笔顺效应是指笔顺不同的笔画在整字识别中具有不同的权重。研究发现,省略首笔画比省略中间或尾部的笔画,对整字识别造成的干扰更大(闫国利等,2013);以笔顺在前的笔画作为启动材料,比笔顺靠后的笔画更有助于汉字命名(Giovanni, 1994)。这些结果也支持笔画是识别过程的基础单元。笔画独立组块效应是指汉字识别的基本单元是单个完整笔画。有研究发现,去除单个完整笔画比去除多个笔画的部分像素产生的干扰更大(Yu et al., 2018)。简单和复杂笔画所含像素数量尽管不同,但识别绩效无差别(张积家等,2002)。以上结果均支持笔画是高于像素的最低计算单元。

既然笔画是汉字产生和识别的起点,那么了解笔画的识别机制,可能是理解汉字识别过程的前提和基础。然而目前尚未见反映笔画表征产生机制的相关证据。笔画识别的本质是,从构成一个汉字的交错线条中分离、抽取与笔画模板相似¹的基本线段的过程。这些线条由书写运动产生,它们交错连接所产生的交接或交叉区域即为节点。相交的数条线,有多种分割线段的方式,所有分割方式构成了一个笔画空间,识别的目的就是在这个空间中寻找最合理的一个分割方案。例如,仅由“横”和“竖钩”两划和一个节点构成的“丁”字,还可能存在两种分割方案:“横折钩”和“短横”;两个“短横”和一个“竖钩”。可见节点附近的所有潜在笔画组合可以由历遍该节点附近每一条线段来产生。对于由 n 个笔画构成的节点,若令其潜在笔画组合的集合为 $A(n)$,则可以推论, $A(n)$ 会随交点相连的笔画数迅速膨胀。同时,对于含有 m 个节点的汉字,若令整字的潜在笔画组合为集合 $A(m)$,则 $A(m)$ 也会随整字包含的节点数增加而迅速增加。可见,识别的主要难点在于笔画空间太大。此外,手写体本身的连笔、缺笔特

性可能造成节点冗余或缺失,使正确识别笔画的难度进一步提高。

对于上述求解空间过大的问题,基于产生式思想的贝叶斯推理算法或许是目前最为有效的求解方法。*Science* 期刊 2015 年 12 月的封面文章介绍了适用于字符识别的贝叶斯规划学习(Bayesian Program Learning, BPL)模型,认为字符的识别是一个基于产生式模型的反向推理过程(Lake et al., 2015)。产生式模型是指,给定目标变量 y 的前提下观测变量 X 的条件概率模型,表示为 $p = (X|Y = y)$ (Meila, 2006)。基于产生式的反向推理则是根据观测变量 X 求解目标变量 Y 的过程,假设有数个目标 $y_1, y_2, \dots, y_n$,如果由产生式模型得出 y_n 产生 X 的概率高于其他目标,即可认为 Y 最有可能由 y_n 产生,则获得了目标变量的最优解。以常见的产生式模型——BPL 模型为例,识别手写字符的过程可以概括如下:先从字母样例中提取节点,枚举所有可能产生该节点的基元(类似于笔画)组合方式。再根据基元间关系的先验知识²,得到各种基元组合的后验概率。最后将输入字符的产生方式和数据库中原型字符的产生方式进行相似度比较,从而推测出哪个字符最有可能“写出”当前观察到的手写字母。该模型在人工字符集的识别任务中达到了人类水平的识别绩效(Lake et al., 2015)。

类比贝叶斯规划学习模型,汉字的产生式识别是对汉字产生过程的逆推理(Yuille & Kersten, 2006)。汉字(独体字)的产生过程可以描述为:首先定义汉字的基本笔画集合,从中抽取数个笔画,并依据先验的笔画组合概率,按照恰当的顺序和空间位置将它们书写出来,从而构成整字图像。产生式识别则是根据整字图像,反推之前的书写过程:从识别输入字的笔画开始,先基于线段交点提取出节点,接着枚举所有能产生该节点的笔画组合方式,再利用笔画先验概率筛选出其中概率最高的组合方案,从而获得输入字的产生方式(Gershman et al., 2012)。当长时记忆中某个原型字的产生方式与之高度相似时,则可将输入字判定为原型字,从而完成汉字识别。就合体字而言,理论上可将合体字视为多个独体字(部件)的嵌套:合体字识别需要先拆分部件,然后执行部件识别过程,从而获得最有可能产生当前输入字的部件组合方案,之后匹配长时记忆中部

¹ 与印刷体汉字不同,手写体笔画变形较大,难以通过与标准笔画直接比对的方式识别。

² BPL 模型认为先验知识包含两类,一是任意两类基元在序列上相邻出现的概率,二是任意两类基元的各种空间关系(尾首相接、首首相接等)的概率。

件产生方式与之最相符的原型字, 进而完成识别³。

此外, 产生式识别依赖的先验知识可能来源于人的书写经验。大量研究表明, 书写能促进汉字识别和词句阅读(朱朝霞 等, 2019): 一方面, 书写动作可以加强汉字正字法相关的视空间表征, 以及形、音、义的联结, 并促进汉字长时动作记忆的形成(Tan et al., 2005); 另一方面, 阅读文字会调用书写相关的空间结构和运动知识, 识别有手写经验的文字相比无经验字会引发感觉运动皮层的更强激活(Cao et al., 2013)。综合产生式理论和经典汉字书写研究, 手写汉字的识别, 可能正是基于节点提供的自下而上信息, 结合由汉字书写习得的自上而下的笔画运动和多层级概率知识, 以产生式的概率推断方法逐步求解正确汉字表征的过程。

假设人以上述过程识别汉字, 则可预测其识别过程应出现以下两种效应: (1)节点数量效应。节点数量是指一个汉字中笔画交叉或相接区域的数量。由于对节点的分析是获得笔画表征的前提, 整字含有的节点越多, 且都得到了充分加工, 则在贝叶斯推测过程中提供的信息量就越大, 即包含较多节点的汉字具有识别优势。(2)节点复杂度效应。对于连结较多笔画的复杂节点, 由于其附近潜在的笔画空间较为庞大, 为了压缩问题空间, 识别过程会更依赖节点的引导作用。因此, 复杂节点为笔画组合的计算提供更大的信息量。本研究通过数个实验, 分别对节点数量效应和节点复杂度效应加以检验。

2 实验 1: 节点数量效应

本实验的目的在于检验节点数量效应, 即包含节点数量较多的汉字具有识别优势。

2.1 被试

实验共招募 26 位浙江大学学生(10 男, 16 女), 年龄范围 17 至 26 岁($M = 21.31$ 岁, $SD = 2.43$ 岁)。所有被试的母语均为汉语且均为右利手。视力或者矫正视力正常。

2.2 实验设计

实验采用 2 (节点数: 多节点和少节点) $\times$ 6 (呈现时间: 10 ms、20 ms、30 ms、40 ms、50 ms、60 ms) 两因素被试内设计。76 个刺激在 6 种呈现时间下分别出现一次, 每个被试共完成 456 个试次。刺激的呈现顺序完全随机。因变量为字判别任务的正确率和反应时。选取多档刺激呈现时间的原因是, 汉字

识别是始于节点分析的时序加工: 识别系统需要先定位节点, 然后拆解附近笔画, 再计算各种笔画组合的概率并取优。因此, 当视觉系统执行到一定加工深度时方能在整字绩效上发现相关效应。由于暂不明确对应的时间窗口, 本实验采用的刺激呈现时间数值参考了采用微观发生法的类似研究(沈模卫, 朱祖祥, 1997; 沈模卫 等, 1998), 因其较好展现了汉字加工变化过程的精细信息。

2.3 实验材料

由于独体字由笔画直接构成, 节点对笔画表征的影响可以较为直接地反映在整字的识别绩效上, 因此本实验使用独体字作为实验材料。实验分为练习阶段和正式实验阶段, 练习阶段的刺激包含 5 个真字和 5 个假字, 正式实验的刺激使用另外 38 个真字和 38 个假字。其中真字又分为多节点组和少节点组, 前者由 19 个包含 3 到 6 个节点的真字组成, 后者由 19 个包含 0 到 2 个节点的真字组成, 多节点组($M = 3.95$, $SD = 0.97$)和少节点组($M = 1.32$, $SD = 0.58$)的节点数差异显著⁴, $df = 1$, $p < 0.001$, $\chi^2 = 29.04$ 。两组材料的笔画数和字频相匹配: 多节点组和少节点组均包括 13 个四笔画字和 6 个五笔画字。通过规模高达两亿字的汉语网络数据库(<http://lingua.mtsu.edu/chinesecomputing/>)确定多节点组的字频范围是 2217~755256 次, 平均为 4.48(转换为以 10 为底的对数), 少节点组的字频是 312~2237915 次, 平均为 4.94(转换为以 10 为底的对数)。假字由真实笔画根据汉字书写习惯组合而成, 也包括多节点和少节点两组, 前者由 19 个含 3 到 6 个节点的假字组成, 后者由 19 个含 0 到 2 个节点的假字组成。多节点组($M = 3.58$, $SD = 0.90$)与少节点组($M = 1.26$, $SD = 0.65$)的节点数差异显著, $df = 1$, $p < 0.001$, $\chi^2 = 29.44$ 。为提升实验的外部效度, 真假字材料都以仿手写字体呈现。使用 Photoshop CC 2018 将所有真假字制作为 150 磅大小的白色字, 放置于边长 130 像素的黑色正方形中心。本实验的真假字样例见图 1。

2.4 实验装置及流程

实验程序采用 Psychtoolbox 编写, 呈现于 17 吋 CRT 屏幕上, 分辨率设为 1024×768 , 刷新率为 100 Hz。实验分为预备实验和正式实验两个阶段。在预备实验阶段, 主试通过展示指导语向被试说明实验的要求和任务。预备实验流程与正式实验一致, 如图 2 所示, 并要求被试进行 20 个试次的练习以熟

³ 产生式框架下, 合体字中节点和笔画的加工过程与独体字一致。此处的部件等高层级加工过程暂是猜想。

⁴ 由于节点数不服从正态分布, 故采用 Mann-Whitney U 检验。



图 1 实验 1 的实验材料示意图

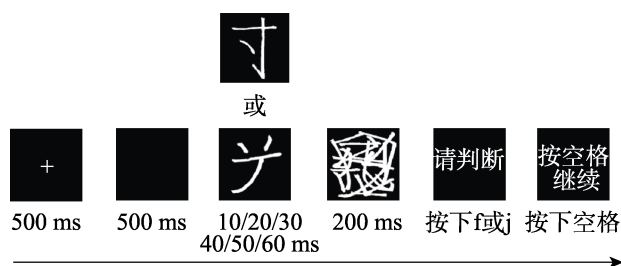


图 2 实验 1 字判别任务示意图

悉实验流程。实验中首先在屏幕中央呈现一个注视点, 短暂空屏后出现一个刺激字。该字可能是真字也可能是假字, 随后加以掩蔽。掩蔽消失后, 要求被试又快又准确地判断刚才的刺激字是真字还是假字, 并按下相应的按键。

2.5 实验结果

本实验中假字组的设置只用于控制被试的反应倾向, 因此只对真字组的正确率, 以及真字组内正确试次和错误试次的反应时加以分析。本文所有实验报告的 F_1 和 F_2 分别是基于被试和基于项目的分析结果。

2.5.1 正确率

正确率的整体平均值为 63%, 标准差为 29%。正确率随呈现时间的变化趋势如图 3a 所示。重复测量方差分析发现, 节点数量的主效应显著, $F_1(1, 25) = 9.65, p = 0.005, \eta_p^2 = 0.28$, 差异的 95% CI = $[-5\%, -1\%]$; $F_2(1, 18) = 4.56, p = 0.047, \eta_p^2 = 0.20$, 差异的 95% CI = $[-5\%, -1\%]$, 多节点字的正确率 ($M = 57\%, SD = 35\%$) 显著高于少节点字 ($M = 55\%, SD = 34\%$)。刺激呈现时间的主效应显著, $F_1(1, 25) = 113.29, p < 0.001, \eta_p^2 = 0.82$; $F_2(3.37, 60.70) = 501.33, p < 0.001, \eta_p^2 = 0.97$, 随呈现时间增长, 正确率显著提高。

基于被试的分析发现两因素间交互作用边缘显著⁵, $F_1(5, 125) = 1.998, p = 0.083, \eta_p^2 = 0.07$; $F_2(5,$

$90) = 1.12, p = 0.355$ 。进一步分析发现: 当刺激呈现时间为 40 ms 时, 被试在真字少节点 ($M = 71\%, SD = 21\%$) 条件下的正确率显著低于真字多节点 ($M = 78\%, SD = 21\%$) 条件, $t(25) = -3.51, p = 0.002$, Cohen's $d = 0.69$, 差异的 95% CI = $[-11\%, -3\%]$ 。在呈现时间为 50 ms 的条件下, 少节点字 ($M = 85\%, SD = 16\%$) 与多节点字 ($M = 88\%, SD = 16\%$) 也存在类似的差异, $t(25) = -2.17, p = 0.040$, Cohen's $d = 0.41$, 差异的 95% CI = $[-7\%, 0\%]$ 。

2.5.2 反应时

反应时的整体平均值为 908 ms, 标准差为 218 ms。不同呈现时间条件下的反应时见图 3b。重复测量方差分析发现, 节点数量的主效应不显著, $F_1(1, 25) = 0.19, p = 0.663$; $F_2(1, 18) = 0.08, p = 0.785$ 。刺激呈现时间的主效应显著⁶, $F_1(1.64, 40.97) = 7.55, p = 0.003, \eta_p^2 = 0.23$; $F_2(5, 90) = 4.01, p = 0.003$, 随呈现时间增长, 反应时显著减少。

基于被试的分析发现两因素间交互作用显著, $F_1(3.75, 93.66) = 3.52, p = 0.012, \eta_p^2 = 0.12$; $F_2(5, 90) = 0.74, p = 0.599$ 。进一步分析发现: 当刺激呈现时间为 40 ms 时, 被试在真字少节点 ($M = 876$ ms, $SD = 160$ ms) 条件下的反应时显著高于真字多节点 ($M = 806$ ms, $SD = 144$ ms) 条件, $t(25) = 3.67, p = 0.001$, Cohen's $d = 0.73$, 差异的 95% CI = $[31$ ms, 110 ms]。在呈现时间为 50 ms 的条件下, 少节点字 ($M = 841$ ms, $SD = 138$ ms) 与多节点字 ($M = 787$ ms, $SD = 135$ ms) 也存在类似的差异, $t(25) = 2.36, p = 0.026$, Cohen's $d = 0.61$, 差异的 95% CI = $[5$ ms, 105 ms]。

2.6 讨论

本实验的结果表明, 节点数量多的字具有识别优势, 表现为更高的正确率和更短的反应时。该效应在刺激呈现时间为 40 ms 和 50 ms 时显著。当刺激呈现时间少于 30 ms 时, 节点数量并未出现上述效应。此时被试对真字的判断正确率也低于随机水平, 说明被试倾向于在无法辨认真假字时做出“假字”的判断。当刺激呈现时间达到 60 ms 时, 被试可以对整字做充分表征, 此时节点数量对整字识别的影响不明显。

综上所述, 在视觉系统能充分获取和分析节点信息的前提下, 节点数量越多, 为笔画拆解过程提供的信息越丰富, 计算系统可以同时利用的有效信

⁵ 由于被试在 40 ms 及以上的呈现时长下方能做出有效真假字判断(正确率高于随机水平 50%), 故独立分析了 40 至 60 ms 的试次, 两因素间交互作用为: $F_1(2, 50) = 4.72, p = 0.013, \eta_p^2 =$

0.16 ; $F_2(2, 36) = 2.48, p = 0.098, \eta_p^2 = 0.12$ 。

⁶ 本文中所有不满足 Mauchly 球形检验的重复测量方差分析结果均采用 Greenhouse-Geisser 方法校正。

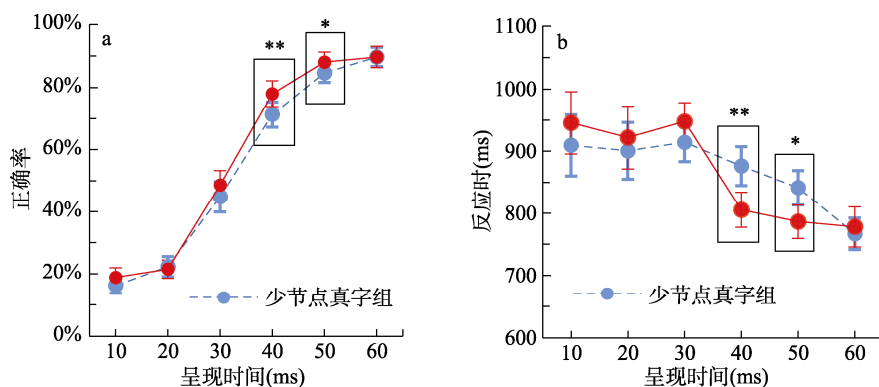


图 3 (a)正确率随呈现时间变化的趋势图 (b)反应时随呈现时间变化的趋势图

注: 误差线代表标准误。* $p < 0.05$, ** $p < 0.01$

息越多, 提高了整字识别的绩效。可见, 本实验验证了人的汉字识别过程具有贝叶斯产生式模型所预测的节点数量效应。

3 实验 2: 节点复杂度效应

本实验的目的在于检验节点复杂度效应, 即连接更多笔画的节点在笔画分离过程中能提供丰富的信息量。

3.1 被试

29 位浙江大学学生(10 男, 19 女)参与正式实验, 年龄范围 18 至 26 岁($M = 21.14$, $SD = 2.18$)。所有被试第一语言均为汉语, 视力或者矫正视力正常。

3.2 实验设计

实验采用 2 (掩盖复杂节点和掩盖简单节点) $\times$ 2 (掩盖第 1 节点和掩盖第 5 节点) $\times$ 4 (呈现时间: 60 ms、70 ms、80 ms、90 ms)三因素被试内设计。160 个刺激在 4 种呈现时间下分别出现一次, 每个被试共完成 640 个试次。刺激的呈现顺序完全随机。因变量为字判别任务的正确率和反应时。

3.3 实验材料

为进一步确认节点在汉字识别中的作用, 本实验使用单个节点被掩盖的合体字作为实验材料。合体字的单个部件可以视为一个独体字, 节点通过干扰部件内的笔画表征, 进而影响部件以及合体字整字的识别。实验分为练习阶段和正式实验阶段, 练习阶段的刺激是 5 个真字和 5 个假字, 正式实验的刺激使用另外的 80 个真字和 80 个假字。

由于笔顺位置不同的笔画在整字识别中具有不同的权重(Giovanni, 1994; 闫国利 等, 2013), 掩盖不同位置的节点对整字识别的影响, 可能会受节点所在笔画的笔顺位置干扰。因此, 本实验将在控制节点顺序的前提下验证节点复杂度效应。节点顺

序为一个汉字按照标准笔顺书写时节点依次产生的顺序。

定义节点的复杂度为构成该节点的笔画数量。本实验中, 由两个笔画形成的节点称为简单节点, 例如“下”中横与竖的交点。有 3 或 4 个笔画形成的节点称为复杂节点, 例如“木”中横、竖、撇、捺的交点。正式实验材料分为真字组和假字组, 真字组分别包含 20 个第 1 节点为简单节点($M = 2.00$, $SD = 0$)和第 1 节点为复杂节点($M = 3.60$, $SD = 0.50$)的真字, 两类字中第 1 节点的复杂度差异显著, $df = 1$, $p < 0.001$, $\chi^2 = 34.82$; 20 个第 5 节点是简单节点($M = 2.00$, $SD = 0$)和 20 个第 5 节点是复杂节点($M = 3.45$, $SD = 0.51$)的真字, 其节点复杂度差异也显著, $df = 1$, $p < 0.001$, $\chi^2 = 34.71$ 。分别掩盖真字的第 1 节点和第 5 节点, 以分离不同节点顺序的效应。假字组对应真字也分为 4 类, 每个假字由对应的真字替换一个部件制得, 且替换的部件不含被掩盖的节点。真字 4 类字组在节点数、笔画数、部件数、整字构型和字频上均加以匹配(见表 1), 呈现材料均使用手写字体。本实验所使用字样如图 4 所示。

3.4 实验装置及流程

实验装置与实验 1 相同。由于实验 1 中刺激呈现时间达到 40 ms 时节点数量才会对被试判断产生影响, 并且本实验的材料是更为复杂的合体字, 因此刺激的呈现时间改为 60 ms 至 90 ms, 其他条件与实验 1 一致。

表 1 实验 2 材料的各种属性

条件	第 1 点复杂	第 5 点复杂	第 1 点简单	第 5 点简单
笔画数	10.85	10.70	10.25	10.30
字频(lg)	4.16	4.35	4.19	4.31
部件数	2.45	2.65	2.55	2.45
节点数	5.75	5.75	5.80	6.00
整字结构	均包含 6 个上下结构合体字和 14 个左右结构合体字			



图 4 实验 2 的实验材料示意图

3.5 实验结果

本实验对真字组的正确率, 以及真字组内正确试次和错误试次的反应时加以分析。

3.5.1 正确率

正确率的整体平均值为 79%, 标准差为 21%。分别掩盖第 1 和第 5 节点时, 字判别任务正确率随呈现时间的变化趋势如图 5 所示。重复测量方差分析发现, 节点复杂度主效应显著⁷, $F_1(1, 28) = 6.93$, $p = 0.014$, $\eta_p^2 = 0.20$, 差异的 95% CI = [-4%, -1%]; $F_2(1, 19) = 0.73$, $p = 0.404$ 。掩盖复杂节点的正确率 ($M = 81\%$, $SD = 21\%$) 显著低于掩盖简单节点的正确率 ($M = 83\%$, $SD = 22\%$)。呈现时间的主效应显著, $F_1(2.22, 62.14) = 8.06$, $p = 0.001$, $\eta_p^2 = 0.22$; $F_2(3, 57) = 7.69$, $p < 0.001$, $\eta_p^2 = 0.29$, 随呈现时间增长, 正确率显著提高。节点顺序的主效应不显著, $F_1(1, 28) = 1.74$, $p = 0.197$; $F_2(1, 19) = 0.32$, $p = 0.580$ 。

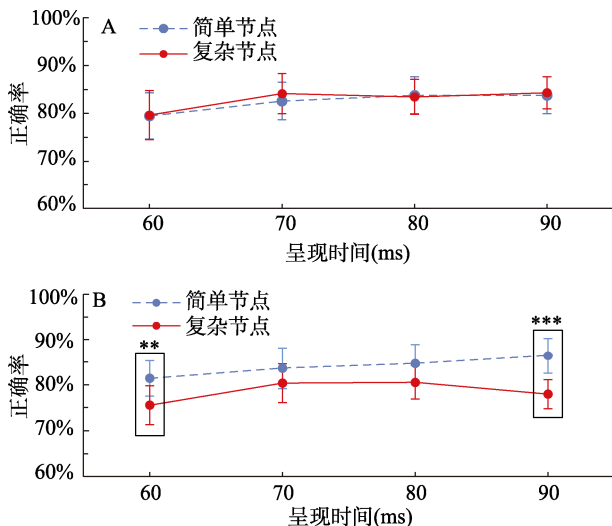


图 5 (A)掩盖第 1 节点时正确率随呈现时间变化的趋势图; (B)掩盖第 5 节点时正确率随呈现时间变化的趋势图

注: 误差线代表标准误。** $p < 0.01$, *** $p < 0.001$ 。

复杂度和节点顺序之间的交互作用显著, $F_1(1, 28) = 11.56$, $p = 0.002$, $\eta_p^2 = 0.29$; $F_2(1, 19) = 2.49$, $p = 0.131$ 。其余的交互作用均不显著: 复杂度和呈现时间之间的交互作用, $F_1(2.29, 64.18) = 1.34$, $p = 0.266$; $F_2(3, 57) = 1.32$, $p = 0.277$ 。节点顺序和呈现时间之间的交互作用, $F_1(3, 84) = 0.16$, $p = 0.926$; $F_2(3, 57) = 0.10$, $p = 0.961$ 。三因素之间的交互作用, $F_1(3, 84) = 0.93$, $p = 0.430$; $F_2(3, 57) = 0.57$, $p = 0.636$ 。

简单效应分析表明, 对于顺序第 5 的节点, 掩盖复杂节点的正确率 ($M = 79\%$, $SD = 18\%$) 显著低于掩盖简单节点 ($M = 84\%$, $SD = 21\%$), $t(116) = -5.22$, $p < 0.001$, Cohen's $d = 0.48$, 差异的 95% CI = [-7%, -3%]。对于顺序第 1 的节点, 掩盖复杂节点的正确率 ($M = 83\%$, $SD = 19\%$) 与掩盖简单节点 ($M = 82\%$, $SD = 19\%$) 无显著差异, $t(116) = 0.57$, $p = 0.572$ 。

3.5.2 反应时

反应时的整体平均值为 929 ms, 标准差为 210 ms。分别掩盖不同位置节点时的反应时见图 6。重复测量方差分析发现, 节点顺序和呈现时间之间的交互作用显著, $F_1(3, 84) = 4.14$, $p = 0.009$, $\eta_p^2 = 0.13$; $F_2(3, 57) = 0.16$, $p = 0.020$, $\eta_p^2 = 0.16$ 。其余的主效应和交互作用均不显著: 节点复杂度的主效应, $F_1(1, 28) = 0.35$, $p = 0.558$; $F_2(1, 19) = 0.33$, $p = 0.575$ 。呈现时间的主效应, $F_1(2.45, 68.56) = 2.27$, $p = 0.100$; $F_2(3, 57) = 2.70$, $p = 0.054$ 。节点顺序的主效应, $F_1(1, 28) = 0.34$, $p = 0.566$; $F_2(1, 19) = 0.08$, $p = 0.780$ 。复杂度和节点顺序之间的交互作用, $F_1(1, 28) = 0.85$, $p = 0.366$; $F_2(1, 19) = 0.16$, $p = 0.693$ 。复杂度和呈现时间之间的交互作用, $F_1(3, 84) = 0.36$, $p = 0.786$; $F_2(3, 57) = 0.52$, $p = 0.668$ 。三因素之间的交互作用, $F_1(3, 84) = 1.05$, $p = 0.376$; $F_2(3, 57) = 1.45$, $p = 0.239$ 。

3.6 讨论

本实验发现, 掩盖复杂节点比简单节点对整字识别产生的干扰更大, 即识别正确率更低。该效应

⁷ 本实验中项目分析不显著, 可能是由于没有控制部件的笔画数、频率、结合律等无关变量。实验 3 对上述因素加以平衡后, 观察到了 F_2 上的显著差异。

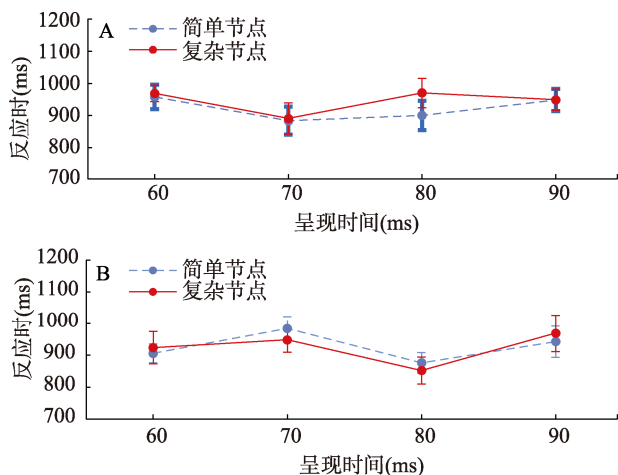


图 6 (A)掩盖第 1 节点时反应时随呈现时间变化的趋势图, (B)掩盖第 5 节点时反应时随呈现时间变化的趋势图

注: 误差线代表标准误。

在掩盖顺序靠后的节点时更为显著。说明高复杂度的节点为笔画拆解过程提供了更为丰富的信息。掩盖处于更大的笔画空间中的复杂节点, 穷举出的笔画组合方式会多于简单节点, 更多的计算量将损害整字的识别绩效。这为贝叶斯产生式模型所预测的节点复杂度效应提供了初步证据。

节点复杂度效应在第 1 节点上比较弱, 可能存在两方面原因: 一是部件的特性削弱了节点复杂度的影响。合体字的第 1 节点通常位于部首上, 部首具有多为形旁、构字能力较强、笔画相对其他部件更少等特性, 其在整字识别中的权重较低, 掩盖此处的节点对整字识别的干扰有限。二是节点的产生方式影响了节点的复杂度。节点由笔画交叉或相接所产生, 在笔画数量一定的前提下, 笔画交叉所产生的节点区域(如“十”字的节点), 相比笔画相交但不穿过的区域(如“厂”字的节点), 前者潜在的笔画组合方式更多, 掩盖该类节点会对笔画拆解和整字识别产生更大干扰。

4 实验 3: 节点复杂度效应的再检验

本实验目的是进一步检验节点复杂度效应。由于声旁和形旁在合体字识别中的作用具有特异性(Lee et al., 2006), 且节点产生方式可能会影响节点复杂程度, 因此本实验将部件类型和节点产生方式作为两个自变量进行操纵。

4.1 被试

26 位浙江大学学生(10 男, 19 女)参与正式实验, 年龄范围 18 至 25 岁($M = 21.50$, $SD = 2.02$)。所有被试第一语言均为汉语, 视力或者矫正视力正常。

4.2 实验设计

实验采用 2 (掩盖复杂节点和掩盖简单节点) $\times$ 2 (掩盖声旁上的节点和掩盖形旁上的节点) $\times$ 2 (节点由笔画交叉所产生和节点由笔画相接但不穿过所产生)三因素被试内设计。240 个刺激在 60 ms 呈现时间下均出现一次, 每个被试共完成 240 个试次。刺激的呈现顺序完全随机。因变量为字判别任务的正确率和反应时。

4.3 实验材料

本实验以单个节点被掩盖的合体字作为实验材料。实验分为练习阶段和正式实验阶段, 练习阶段的刺激是 5 个真字和 5 个假字, 正式实验的刺激使用另外的 120 个真字和 120 个假字。

简单节点和复杂节点的定义同实验 2。正式实验材料包含 120 个真字刺激和 120 个假字刺激。真字均为左形右声结构的形声字(符合实验要求的左声右形汉字数量稀少, 故不采用), 分为 8 组, 每组 15 个目标刺激。假字组对应真字也分为 8 组, 每个假字由对应的真字替换一个部件制得, 且替换的部件不含被掩盖的节点。8 组真字在笔画数、字频、部件数、形旁笔画数、形旁频率、形旁结合律、声旁笔画数、声旁频率和声旁结合律上均加以匹配(见表 2), 真假字材料均使用手写字体。本实验所使用字样如图 7 所示。

4.4 实验装置及流程

实验装置与实验 2 相同。由于实验 2 中刺激呈现时间达到 60 ms 时已可观察到节点复杂度效应, 因此本实验中刺激的呈现时间均采用 60 ms, 其他条件与实验 2 一致。

4.5 实验结果

本实验对真字组的正确率, 以及真字组内正确试次和错误试次的反应时加以分析。

4.5.1 正确率

数据的整体平均值为 72%, 标准差为 18%。字判别任务正确率如图 8 所示。方差分析发现, 节点复杂度的主效应显著, $F_1(1, 200) = 8.32$, $p = 0.004$, $\eta_p^2 = 0.04$; $F_2(1, 112) = 5.69$, $p = 0.019$, $\eta_p^2 = 0.05$, 掩盖复杂节点的正确率($M = 68\%$, $SD = 19\%$)显著低于掩盖简单节点的正确率($M = 76\%$, $SD = 16\%$)。部件类型的主效应不显著, $F_1(1, 200) = 0.01$, $p = 0.911$; $F_2(1, 112) = 0.02$, $p = 0.902$ 。节点产生方式的主效应不显著, $F_1(1, 200) = 1.70$, $p = 0.193$; $F_2(1, 112) = 1.42$, $p = 0.236$ 。

基于被试的分析发现, 复杂度和节点产生方式

表 2 实验 3 材料的各种属性

条件	CIP	CJP	CIS	CJS	SIP	SJP	SIS	SJS
整字笔画数	11.27	11.47	11.27	10.80	11.67	11.47	11.00	11.73
整字频率(lg)	2.77	2.93	2.76	2.69	2.81	2.68	2.86	2.92
形旁笔画数	4.87	4.60	4.60	4.73	4.60	4.13	4.13	4.57
形旁频率(lg)	4.91	5.00	4.62	4.95	5.03	5.02	5.06	4.84
形旁结合律(lg)	2.38	2.44	2.37	2.09	2.28	2.18	2.32	2.13
声旁笔画数	6.40	6.87	6.93	6.73	7.06	6.73	6.87	7.27
声旁频率(lg)	4.90	4.66	5.03	4.93	4.67	4.84	4.93	4.85
声旁结合率(lg)	1.98	1.87	1.95	1.73	1.68	1.72	2.00	1.89

注: C 和 S 分别表示“复杂节点”和“简单节点”, I 和 J 分别表示“由笔画交叉产生”和“由笔画相接产生”, P 和 S 分别表示“位于声旁”和“位于形旁”。



图 7 实验 3 的实验材料示意图

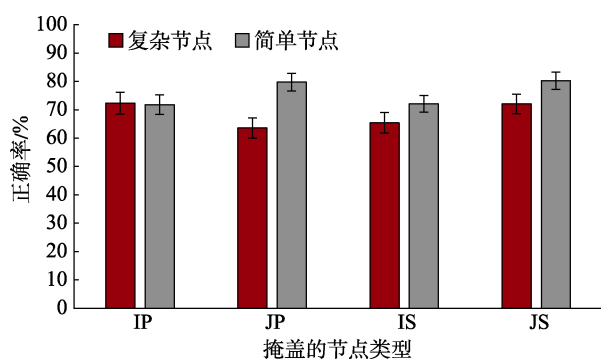


图 8 字判别任务的正确率

注: 误差线代表标准误。IP 表示笔画交叉产生的形旁节点, IS 表示笔画交叉产生的声旁节点, JP 表示笔画相接产生的形旁节点, JS 表示笔画相接产生的声旁节点。

的交互作用显著⁸, $F_1(1, 200) = 3.87, p = 0.050, \eta_p^2 = 0.02$; $F_2(1, 112) = 1.96, p = 0.165, \eta_p^2 = 0.02$ 。复杂度和部件类型之间的交互作用不显著, $F_1(1, 200) = 0.01, p = 0.908$; $F_2(1, 112) = 0.00, p = 1.000$ 。部件类型和节点产生方式之间的交互作用不显著, $F_1(1, 200) = 0.08, p = 0.092$; $F_2(1, 112) = 1.42, p = 0.236$ 。三因素之间的交互作用亦不显著, $F_1(1, 200) = 3.24, p = 0.073$; $F_2(1, 112) = 1.326, p = 0.252$ 。

进一步分析表明, 对于笔画相接所产生的节点,

掩盖复杂节点的正确率($M = 68\%, SD = 18\%$)显著低于掩盖简单节点($M = 80\%, SD = 16\%$), $t(102) = -3.67, p < 0.001$, Cohen's $d = 0.73$, 差异的 95% CI = $[-19\%, -6\%]$ 。对于笔画交叉产生的节点, 掩盖复杂节点的正确率($M = 69\%, SD = 19\%$)与掩盖简单节点($M = 72\%, SD = 16\%$)无显著差异, $t(102) = -0.89, p = 0.378$ 。

4.5.2 反应时

数据的整体平均值为 1240 ms, 标准差为 386 ms。字判别任务反应时见图 9。方差分析未发现任何主效应(节点复杂度, $F_1(1, 200) = 0.17, p = 0.678$; $F_2(1, 112) = 0.93, p = 0.338$ 。部件类型, $F_1(1, 112) = 0.21, p = 0.651$; $F_2(1, 112) = 0.86, p = 0.357$ 。节点

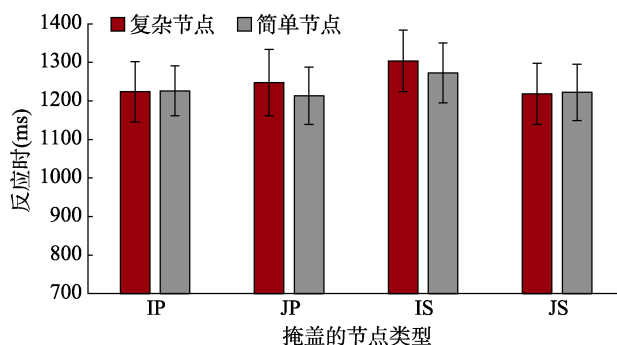


图 9 字判别任务的反应时

注: 误差线代表标准误。IP 表示笔画交叉产生的形旁节点, JP 表示笔画相接产生的形旁节点, IS 表示笔画交叉产生的声旁节点, JS 表示笔画相接产生的声旁节点。

⁸ 鉴于节点附近笔画识别机制的研究刚起步, 在部件和笔画层次的常规控制变量之外, 可能存在未经控制的与节点产生方式有关的未知变量, 致使项目分析未能显著。将来的工作应寻找影响笔画拆分的关键因素。

产生方式, $F_1(1, 200) = 0.40, p = 0.527$; $F_2(1, 112) = 1.45, p = 0.231$ 。

未发现任何交互作用(节点复杂度和部件类型之间的交互作用不显著, $F_1(1, 200) = 0.01, p = 0.909$; $F_2(1, 112) = 0.18, p = 0.676$ 。节点复杂度和节点产生方式的交互作用不显著, $F_1(1, 200) = 0.06, p = 0.804$; $F_2(1, 112) = 0.52, p = 0.474$ 。部件类型和节点产生方式的交互作用不显著, $F_1(1, 200) = 0.85, p = 0.356$; $F_2(1, 112) = 3.13, p = 0.08$ 。三因素之间的交互作用亦不显著, $F_1(1, 200) = 0.10, p = 0.751$; $F_2(1, 112) = 0.75, p = 0.390$ 。

从差异方向上看, 反应时模式与正确率是基本一致的(正确率高的反应时短), 不存在反应时-正确率权衡。

4.6 讨论

本实验发现掩盖复杂节点比简单节点对整字识别产生的干扰更大, 即识别正确率更低。该效应在掩盖笔画相接但不穿过所产生的节点时更为显著。

节点复杂度和节点产生方式存在交互作用。当掩盖笔画交叉产生的节点时, 简单和复杂节点的正确率均较低, 可能是交叉这一几何形式扩充了潜在的笔画组合方案, 增加了掩盖简单节点时认知计算的难度。该交互作用的项目分析结果并不显著, 表明可能存在未经控制的与节点产生方式有关的未知变量, 提示下一步研究可探寻笔画拆分的关键影响因素; 节点复杂度和部件类型不存在交互作用, 可能是由于实验材料均为左形右声的形声字, 掩盖形旁和声旁节点分别意味着干扰顺序在前和在后笔画的表征, 即笔画顺序效应和声旁优势效应相互抵消。以往研究为此提供了佐证: 相比形旁, 形声字的声旁在整字识别中的作用更为明显; 而相比顺序靠后的笔画, 顺序在前的笔画对整字识别更为重要(闫国利 等, 2013)。因此, 两种因素相平衡后, 节点复杂度在声旁和形旁上的影响没有明显差异。

本实验通过更加严格的实验控制, 进一步说明高复杂度的节点提供了更为丰富的笔画拆解信息, 掩盖处于更大的笔画空间中的复杂节点会显著增加认知过程的计算难度。可见, 人的汉字识别过程具有贝叶斯产生式模型所预测的节点复杂度效应。

5 总讨论

本研究基于字符识别的贝叶斯规划学习模型, 认为汉字字形识别是一个产生式的反向推理过程, 提出并验证了节点数效应和节点复杂度效应的预

测。三个实验说明节点提供给贝叶斯推断过程的信息量越多, 整字越容易识别。以上结果为汉字字形识别产生式过程提供了证据, 表明对节点的加工是识别过程的基础。

5.1 节点分析是获得笔画表征的前提

本研究的结果表明, 笔画表征的获得依赖于节点提供的信息量增益。实验 1 发现, 整字包含的节点数量越多, 识别绩效越好, 从数量的角度表明节点为笔画分割提供了引导信息。实验 2、3 发现, 掩盖的节点越复杂, 整字识别绩效越差, 从性质的角度表明不同类型的节点提供的引导信息存在差异。汉字字形是二维平面中线条的集合, 获得笔画表征需要经历从线条中分离、抽取的过程, 且不存在唯一解。因此, 节点为获得恰当合理的笔画表征提供了自下而上的笔画分离线索。

节点为笔画分割过程提供的引导信息, 本质上可能是节点蕴含的笔画运动信息, 即有关该节点和邻近笔画是如何由汉字书写所“产生”的。笔画是由自左向右、自上而下两条运动规则产生的单向线段。没有线段交错的部分, 笔画的产生方式几乎是确定的。只有在线段交错的节点处, 笔画在不违背产生规则的前提下具有多种可能的运动方向。因此分割线段的重点是在节点处, 此处的笔画运动信息最为丰富。这与经成分分析理论认为笔画曲折、交接的地方具有更多的非偶然性特征相一致(Huang & Wang, 1992)。笔画运动信息经由汉字书写训练习得, 先前大量研究发现书写能力与阅读能力存在正相关关系(朱朝霞 等, 2019)。认知神经科学的证据表明, 两者共享左侧梭状回和左侧额下回等神经网络, 阅读过程中书写相关的运动功能区也会激活。行为学研究也发现, 汉字书写中的运动规划能够促进汉字的长时动作记忆的形成(Tan et al., 2005), 人们可以在书写汉字的过程中学习到笔画运动方向的知识, 该类知识会自上而下影响对笔画的识别(Tse & Cavanagh, 2000)。因此, 节点蕴含的笔画运动信息是视觉系统能从静态整字图像中拆分出笔画组合的重要线索之一。

就汉字加工的时间进程而言, 节点的表征与分析可能是汉字识别的初始环节。实验 1 发现, 在刺激呈现时间较短(小于 40 ms)时, 节点信息尚未充分提取, 即多节点字和少节点字的识别绩效未表现出差别, 整字识别的正确率并未超出随机水平(50%)。呈现时间达到 40 ms 后, 节点获得较好表征, 整字识别正确率开始超过随机水平。可见, 整字能被有

效识别时对应的刺激呈现时长与节点数量效应出现时的呈现时长一致。因此,节点可能是整字识别刺激输入后较早获得的基础特征,利用节点信息后才能执行笔画分离过程,从而有效识别整字。

本研究支持笔画分割以并行加工的方式进行。实验 1 发现识别多节点字的识别绩效更好,表明节点越多,视觉系统可以同时利用的节点也越多。节点多带来的信息量增益可以在不消耗更多时间的前提下使识别的后验概率更快达到阈值,从而表现出更好的整字识别绩效。这种并行加工的特性,高效利用了节点提供的笔画分割信息,从而更容易找到包容所有节点的整字产生方式。这符合前人发现的视觉系统可以对单一刺激维度的多个项目做并行加工的特性(Cave & Wolfe, 1990; Treisman & Gelade, 1980; Treisman, 1982)。该特性使视觉系统在整字范围内搜索和提取线段交错点,进而利用冗余节点提供的信息促进整字识别。图形识别的相关研究也支持冗余信息量的促进效应:中等复杂度的图形因其具有相对更高的冗余度,识别速度比简单图形更快(Lockhead & Pomerantz, 1991)。

5.2 节点是客体识别的通用特征

本研究发现,节点是视觉系统加工汉字刺激的重要特征,掩盖节点会对合体字的识别产生干扰。先前曾有汉字节点的相关研究发现,掩盖节点后整字识别的正确率降低,反应时增加,提示汉字节点可能是类似物体轮廓线交点的关键特征(骆非凡, 2020)。然而该研究未涉及节点的信息量内涵。拼音文字相关研究的结果与之类似:有实验采用掩蔽启动范式,发现包含笔画节点特征的启动刺激会促进被试对英文字母的命名(Petit & Grainger, 2002);另有实验发现,去除英文字母中线段的节点比掩盖线段中间部分对字母命名造成的干扰更大(Lanthier et al., 2009)。一些研究认为,图像中轮廓线条的节点不单在字符识别中起作用,对客体识别均有重要意义(Dehaene et al., 2005; Dehaene, 2009)。例如,掩盖物体轮廓线的交点比掩盖轮廓线的中点对识别的干扰更大(Biederman, 1987),保留轮廓线交点的图形命名正确率更高(Szwed et al., 2011)。以上证据表明,节点在客体识别中广泛发挥作用,这意味着节点可能是客体识别的一种通用特征。

节点在字符识别中的作用可能来源于视觉系统早期已具备的客体识别机制。神经回路回收假设(Neuronal recycling hypothesis)认为,在进化早期,人类的文字阅读能力并不存在先天的专门功能区,

而是重塑功能较为适合的、但原本用于其他功能的脑区,使之适应文字阅读这一新功能(Dehaene et al., 2005)。节点之所以在字符识别中起重要作用,是源于人类对物体的识别依赖于节点。该假说获得了实证研究的支持:有研究分别以掩盖和保留线段节点的字母、物体作为实验材料,发现被试观看时保留节点的材料时梭状回激活程度更高(Szwed et al., 2011);另有研究发现,恒河猴颞下皮层的部分神经元会对含有线段节点的图形产生明显的响应(BrinCAT & Connor, 2004)。此外,从文字符号产生的历史看,大多数文明使用的文字符号均是线条的排布组合(Changizi et al., 2006)。这些文字系统以这一形式诞生,可能是视觉系统已经具备了编码这些图形的能力,因此选择这些图形来创造文字系统。总之,汉字识别中的节点特征加工机制可能源于客体识别的相应机制。

5.3 本研究对汉字字形识别过程的启示

本研究验证了节点数量效应和节点复杂度效应,其结果仅为产生式识别过程的早期阶段提供了证据。识别系统后续会利用先验的笔画关系知识,从多种笔画拆分方案中推断哪一种最可能产生当前的输入字。根据贝叶斯规划学习模型的计算特性预测,上述识别过程还会表现出笔画概率自主学习效应和高频笔画组合优势效应。

笔画概率自主学习效应即实现概率推断的前提是识别系统具备笔画概率自主学习模块,笔画关系概率知识可以通过对字符样例的统计学习自主掌握。一些研究为上述预测提供了支持:有实验发现整字中同一笔画所占的比例可以作为正字法知识通过内隐学习获得(王菲, 2015);另有研究在高强度练习和正确反馈条件下,发现在完成笔画维度特征分类任务时内隐学习有效(侯伟康, 奏启庚, 1996)。以上结果说明存在专门学习笔画相关概率的认知模块。

高频笔画组合优势效应是指由于高频率的笔画组合具有较高的先验概率,识别系统面对含有高频笔画组合的字有计算优势。已有大量研究支持汉字家族的促进效应,形旁构字能力大的汉字更容易被辨认(张积家, 姜敏敏, 2008; Su & Weekes, 2007),特别是促进高频字的识别,同时抑制低频字的识别(钱怡 等, 2015)。上述结果说明高频的笔画组合具有促进效应。可见,现有研究一定程度为产生式识别过程的后期阶段提供了佐证。

本研究也表明,合体字的字形识别建立在独体

字识别的产生式过程之上, 笔画是合体字加工的层次之一。一些研究认为部件是合体字识别的基本单元, 其作用主要表现在部件数、部件频率、部件位置、部件类型等维度(韩布新, 1998; 张积家, 姜敏敏, 2008; Chen & Yeh, 2017)。实验2和3发现掩盖合体字中复杂节点比掩盖简单节点对整字识别产生的干扰更大, 说明节点在合体字识别中同样发挥作用, 经节点拆分出的笔画也是合体字识别所需的表征单元。上述结果支持笔画和部件均是合体字加工单元的理论, 且符合大多数经典模型(罗艳琳 等, 2008; 彭聃龄, 王春茂, 1997; 闫国利 等, 2013; Taft & Zhu, 1997)。近期有研究发现, 部件间具有层级关系(张瑞, 2017), 合体字的识别可能是先依据独体字识别的产生式过程识别出浅层级的部件, 然后以同样的过程识别出深层级的部件(沈模卫 等, 1997, 1998)。本研究主要关注笔画拆分过程, 未能阐释部件如何参与识别过程, 未来可以依托产生式思想展开探索, 以建立系统的汉字字形识别模型。

作为模拟人类认知过程的计算模型, 字形识别产生式过程反映了人类智能拥有组成性、因果关系和自学习三种特性(Lake et al., 2015), 分别对应于: (1)计算主体对汉字的表征具有层级结构。笔画构成了整字, 整字可以由节点拆解为笔画; (2)笔画之间存在因果联系。两个连续笔画之间具有共同出现的概率和具备特定空间关系的概率, 前一笔画不仅规定了后一笔画可能是什么, 也规定了后一笔画可能的空间位置, 后一笔是前一笔的“果”; (3)计算主体可以从字符中学习到笔画关系的概率分布, 并根据观察到的新样本更新已有先验。这三种特性使人类可以实现基于少量样本的学习, 大幅压缩计算空间, 从而表现出人类智能相对于当前人工智能的优越性(唐宁 等, 2018)。

6 小结

本研究发现汉字包含的节点数量越多, 整字越容易识别, 且被掩盖的节点越复杂, 对整字识别的干扰越大。说明视觉系统以并行加工方式按照线段节点分离笔画, 笔画表征的获得依赖于节点提供的笔画分离信息, 信息越丰富整字识别绩效越好。研究增进了对汉字字形识别早期视觉过程的认识, 为字形识别产生式过程提供了证据。

参 考 文 献

Biederman, I. (1987). Recognition-by-components: a theory of

- human image understanding. *Psychological Review*, 94(2), 115-47.
- Brincat, S. L., & Connor, C. E. (2004). Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nature Neuroscience*, 7(8), 880-886.
- Cao, F., Vu, M., Chan, D. H. L., Lawrence, J. M., Harris, L. N., Guan, Q., ... Perfetti, C. A. (2013). Writing affects the brain network of reading in Chinese: A functional magnetic resonance imaging study. *Human Brain Mapping*, 34(7), 1670-1684.
- Cave, K. R., & Wolfe, J. M. (1990). Modeling the role of parallel processing in visual search. *Cognitive Psychology*, 22(2), 225-271.
- Changizi, M. A., Zhang, Q., Ye, H., & Shimojo, S. (2006). The structures of letters and symbols throughout human history are selected to match those found in objects in natural scenes. *American Naturalist*, 167(5), E117-E139.
- Chen, C. F., & Huang, X. T. (2004). *Structural symmetry of Chinese character cognition: Research and application*. Beijing: Xinhua Publishing House.
- [陈传锋, 黄希庭. (2004). 结构对称性汉字认知: 研究与应用. 北京: 新华出版社.]
- Chen, T. Q. (1983). *Chinese orthography*. Wuhan, China: Hubei People's Press.
- [陈天泉. (1983). 汉字正字法. 武汉: 湖北人民出版社.]
- Chen, Y. C., & Yeh, S. L. (2017). Examining radical position and function in Chinese character recognition using the repetition blindness paradigm. *Language Cognition & Neuroscience*, 32(1), 37-54.
- Dehaene, S. (2009). *Reading in the brain*. New York, America: Penguin Press.
- Dehaene, S., Duhamel, J.-R., Hauser, M. D., & Rizzolatti, G. (2005). *From monkey brain to human brain*. Cambridge, America: MIT Press.
- Gershman, J., Vul, E., & Tenenbaum, B. (2012). Multistability and perceptual inference. *Neural Computation*, 24(1), 1-24.
- Giovanni, F. B. D. (1994). Order of strokes writing as a cue for retrieval in reading Chinese characters. *European Journal of Cognitive Psychology*, 6(4), 337-355.
- Grenander, U. (1976). *Pattern synthesis: Lectures in pattern theory I*. New York, Springer-Verlag.
- Han, B. X. (1998). The frequency effect of character constituents in the recognition of Chinese characters. *Journal of Psychological Science*, 21(3), 193-195+286.
- [韩布新. (1998). 汉字识别中部件的频率效应. 心理科学, 21(3), 193-195+286.]
- Hou, W. K., & Qin, Q. G. (1996). A preliminary study of the implicit learning of Chinese characters. *Journal of Psychological Science*, 19(6), 351-354+384.
- [侯伟康, 秦启庚. (1996). 汉字特征内隐学习的初步实验研究. 心理科学, 19(6), 351-354+384.]
- Huang, J. T., & Wang, M. Y. (1992). From unit to gestalt: perceptual dynamics in recognizing Chinese characters. *Advances in Psychology*, 90, 3-35.
- Kok, P., Brouwer, G., van Gerven, M., & de Lange, F. (2013). Prior expectations bias sensory representations in visual cortex. *Journal of Neuroscience*, 33(41), 16275-16284.
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25(2), 1097-1105.
- Laeng, B., Bloem, I. M., D'Ascenzo, S., & Tommasi, L. (2014). Scrutinizing visual images: The role of gaze in mental imagery and memory. *Cognition*, 131(2), 263-283.
- Lake, B. M., Salakhutdinov, R., & Tenenbaum, J. B. (2015).

- Human-level concept learning through probabilistic program induction. *Science*, 350(6266), 1332–1338.
- Lanther, S. N., Risko, E. F., Stolz, J. A., & Besner, D. (2009). Not all visual features are created equal: Early processing in letter and word recognition. *Psychonomic Bulletin & Review*, 16(1), 67–73.
- Lee, C. Y., Tsai, J. L., Huang, H. W., Hung, D. L., & Tzeng, O. J. L. (2006). The temporal signatures of semantic and phonological activations for Chinese sublexical processing: An event-related potential study. *Brain Research*, 1121(1), 150–159.
- Li, Z., Wu, Q., Xiao, Y., Jin, M., & Lu, H. (2020). Deep matching network for handwritten Chinese character recognition. *Pattern Recognition*, 107, 107471.
- Lockhead, G. R., & Pomerantz, J. R. (1991). *The perception of structure*. Washington, DC, America: American psychological Association.
- Luo, F. F. (2020). *Potent features for identification of Chinese character strokes* (Unpublished master's thesis). Hangzhou Normal University, China.
- [骆非凡. (2020). 汉字笔画认知的关键性特征 (硕士学位论文文). 杭州师范大学.]
- Luo, Y. L., Chen, M., & Peng, D. L. (2008). The influence of the whole words, radicals and strokes in stroke-form-character. *Psychological Exploration*, 28(1), 28–32.
- [罗艳琳, 陈墨, 彭聘龄. (2008). 整字、部件与笔画对笔画构成汉字的影响. *心理学探新*, 28(1), 28–32.]
- Meila, M. (2006). Machine learning: Discriminative and generative. *The Mathematical Intelligencer*, 28(1), 67–69.
- Peng, D. L., & Wang, C. M. (1997). Basic processing unit of Chinese character recognition: Evidence from stroke number effect and radical number effect. *Acta Psychologica Sinica*, 29(1), 9–17.
- [彭聘龄, 王春茂. (1997). 汉字加工的基本单元: 来自笔画数效应和部件数效应的证据. *心理学报*, 29(1), 9–17.]
- Petit, J. P., & Grainger, J. (2002). Masked partial priming of letter perception. *Visual Cognition*, 9(3), 337–353.
- Qian, Y., Zhang, W. Y., & Bi, H. Y. (2015). The effect of semantic-radical family in Chinese character recognition. *Chinese Journal of Ergonomics*, 21(3), 25–30.
- [钱怡, 张逸玮, 毕鸿燕. (2015). 汉字识别中的形旁家族效应. *人类工效学*, 21(3), 25–30.]
- Ren, X. Q., Fang, X., Sui, X., & Wu, Y. (2018). Characteristics and neural mechanisms of handwritten character recognition. *Advances in Psychological Science*, 26(7), 1174–1185.
- [任晓倩, 方嫻, 隋雪, 吴岩. (2018). 手写体文字识别的特点及神经机制. *心理科学进展*, 26(7), 1174–1185.]
- Shen, M. W., Pan, S. H., & Chen, X. (1998). A study on the cognitive process of identification of two-radical Chinese characters. *Chinese Journal of Applied Psychology*, 4(1), 27–32.
- [沈模卫, 潘善会, 陈新. (1998). 合体汉字字形识别过程探索. *应用心理学*, 4(1), 27–32.]
- Shen, M. W., Pan, S. H., & Li, Z. P. (1997). The effects of global perception of Chinese character morphology on recognition of its radical. *Chinese Journal of Applied Psychology*, 3(1), 47–51.
- [沈模卫, 潘善会, 李忠平. (1997). 整体字形对部件识别的影响. *应用心理学*, 3(1), 47–51.]
- Shen, M. W., & Zhu, Z. X. (1997). Experimental tests of the hierarchical model of two-radical Chinese character perception. *Acta Psychologica Sinica*, 29(4), 350–356.
- [沈模卫, 朱祖祥. (1997). 对汉字字形识别层次模型的实验验证. *心理学报*, 29(4), 350–356.]
- Su, I., & Weekes, B. (2007). Effects of frequency and semantic radical combinability on reading in Chinese: An ERP study. *Brain and Language*, 103(1-2), 111–112.
- Szwed, M., Dehaene, S., Kleinschmidt, A., Eger, E., Valabregue, R., Amadon, A., & Cohen, L. (2011). Specialization for written words over objects in the visual cortex. *Neuroimage*, 56(1), 330–344.
- Taft, M., & Zhu, X. (1997). Submorphemic processing in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 23(3), 761–775.
- Tan, L. H., Spinks, J. A., Eden, G. F., Perfetti, C. A., & Siok, W. T. (2005). Reading depends on writing, in Chinese. *Proceedings of the National Academy of Sciences of the United States of America*, 102(24), 8781–8785.
- Tang, N., An, W., Xu, H. K., Zhou, J. F., Gao, T., & Shen, M. W. (2018). From “data” to “Representation”: How human cognition can inspire artificial intelligence. *Chinese Journal of Applied Psychology*, 24(1), 3–14.
- [唐宁, 安玮, 徐昊骅, 周吉帆, 高涛, 沈模卫. (2018). 从数据到表征: 人类认知对人工智能的启发. *应用心理学*, 24(1), 3–14.]
- Treisman, A. M. (1982). Perceptual grouping and attention in visual search for features and for objects. *Journal of Experimental Psychology*, 8(2), 194–214.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12(1), 97–136.
- Tse, P. U., & Cavanagh, P. (2000). Chinese and Americans see opposite apparent motions in a Chinese character. *Cognition*, 74(3), B27–B32.
- Wang, F. (2015). *The implicit learning research of Chinese words character and form-meaning rules for learning difficulty students* (Unpublished master's thesis). Zhengzhou University, China.
- [王菲. (2015). 学困生汉语字词特征及形义规则的内隐学习研究 (硕士学位论文文). 郑州大学.]
- Yan, G. L., Chi, H., Bian, Q., Xu, Z. J., & Cui, L. (2013). Effects of phonetic radicals and pattern of stroke removal on identification of Chinese phonograms. *Journal of Psychological Science*, 36(1), 14–19.
- [闫国利, 迟慧, 卞迁, 徐子珺, 崔磊. (2013). 声旁与笔画省略方式对形声字识别的影响. *心理科学*, 36(1), 14–19.]
- Yu, L., Xiong, J., Zhang, Q., Drieghe, D., & Reichle, E. D. (2018). Eye-movement evidence For the mental representation of strokes in Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(3), 544–551.
- Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: Analysis by synthesis? *Trends in Cognitive Sciences*, 10(7), 301–308.
- Zhang, J., Du, J., & Dai, L. (2020). Radical analysis network for learning hierarchies of Chinese characters. *Pattern Recognition*, 103, 107305.
- Zhang, J. J., & Jiang, M. M. (2008). The effects of the radical family size and high frequency phonetic radical family member on phonogram recognition. *Acta Psychologica Sinica*, 40(9), 947–960.
- [张积家, 姜敏敏. (2008). 形旁家族、声旁家族和高频同声旁字对形声字识别的影响. *心理学报*, 40(9), 947–960.]
- Zhang, J. J., Wang, H. P., Zhang, M., & Zhang, H. C. (2002). The effect of the complexity and repetition of the strokes on the cognition of the strokes and the Chinese characters. *Acta Psychologica Sinica*, 34(5), 7–11.
- [张积家, 王惠萍, 张萌, 张厚粲. (2002). 笔画复杂性和重复性对笔画和汉字认知的影响. *心理学报*, 34(5), 7–11.]
- Zhang, R. (2017). *The role of hierarchical organization of*

- Chinese character components in vocabulary recognition* (Unpublished master's thesis). Shaanxi Normal University, China.
- [张瑞. (2017). 汉字部件的层级组织在词汇识别中的作用 (硕士学位论文). 陕西师范大学.]
- Zhang, W. T., & Feng, L. (1992). A study on the unit of processing in recognition of Chinese characters. *Acta Psychologica Sinica*, 24(4), 45–51.
- [张武田, 冯玲. (1992). 关于汉字识别加工单位的研究. *心理学报*, 24(4), 45–51.]
- Zheng, C. M., & Gao, S. R. (1982). *Psychological research on Chinese language*. Taipei, China: Wenhe Publishing Co., Ltd.
- [郑昭明, 高尚仁. (1982). *中国语文的心理学研究*. 台北: 文鹤出版有限公司.]
- Zhou, J. F., Xu, H. K., Tang, N., Shi, B. H., Zhao, Y., Gao, T., & Shen, M. W. (2016). The “strong-cognition” psychology research: Implications of AlphaGo. *Chinese Journal of Applied Psychology*, 22(1), 3–11.
- [周吉帆, 徐昊骅, 唐宁, 史博皓, 赵阳, 高涛, 沈模卫. (2016). “强认知”的心理学研究: 来自 alphago 的启示. *应用心理学*, 22(1), 3–11.]
- Zhu, Z. X., Liu, L., Cui, L., & Peng, D. L. (2019). The influence of writing on reading: Evidence from the contrast between traditional writing and typing. *Advances in Psychological Science*, 27(5), 796–803.
- [朱朝霞, 刘丽, 崔磊, 彭聘龄. (2019). 书写对阅读的影响——来自传统书写与电脑打字的证据. *心理科学进展*, 27(5), 796–803.]

The role of stroke nodes in the recognition of handwritten Chinese characters

ZHU Yiming, ZHAO Yang, TANG Ning, ZHOU Jifan, SHEN Mowei

(Department of Psychology and Behavioural Sciences, Zhejiang University, Hangzhou 310058, China)

Abstract

Generative theory holds that the recognition of visual graphics is the inverse reasoning of its generation process. Chinese characters are hieroglyphs formed by interlacing strokes according to orthographic rules. Chinese character recognition can be regarded as the reverse reasoning of the generation process of Chinese characters. Based on the typical generative model -- Bayesian program learning model, the recognition of Chinese characters starts from recognizing the strokes. Firstly, the nodes are extracted based on the intersection of lines, and then all the stroke combination modes that can generate the node are enumerated to obtain the generation mode of Chinese characters. According to the above prediction, the number of nodes and node complexity are important factors in the process of Chinese character recognition. This study investigated the role of nodes in Chinese character recognition through three experiments.

If the nodes provide guidance information for stroke segmentation, the more nodes, the better the performance of Chinese character recognition. In Experiment 1, we tested whether characters with more nodes have recognition advantages by adopting a 2×2 within-subjects design and using 76 single characters as the materials. Characters were chosen from two groups (high node-count and low node-count) of true characters, and two groups (high node-count and low node-count) of fake characters. The characters were briefly presented (10 ms, 20 ms, 30 ms, 40 ms, 50 ms, 60 ms) and appeared once at each presentation time. The presentation order of stimuli was completely random. Each participant completed a total of 456 trials. Twenty-six participants joined in the experiment. After observing each character, the participants reported whether it was a true character or a fake one. If high complex nodes in a larger stroke space provide more information, covering high complex nodes will cause greater interference to character recognition. In Experiment 2, we tested whether characters covered the high complex nodes are harder to recognize by adopting a 2×2×4 within-subjects design and using 160 compound characters covering a node as the materials. Characters were chosen from four groups (covering the first node with high or low complexity and the fifth node with high or low complexity) of true characters, and four groups of fake characters with the same conditions. The process was the same as that of Experiment 1, except the presentation time (60 ms, 70 ms, 80 ms, 90 ms). Twenty-nine participants joined in the experiment. Each participant completed 640 trials. Experiment 3 adopted a task similar to Experiment 2, and added two variables: component type and node generation method. The presentation time was 60 ms. Characters were chosen from eight groups of true characters, and eight groups of fake characters with the same conditions. Each stimulus is presented once. Twenty-six participants joined in the experiment. The accuracy and reaction time (RT) of true characters were analyzed in all experiments.

The results showed that the participants had a better recognition performance for the characters with more nodes (node number effect), and covering the high complex nodes significantly damaged their performance (node complexity effect). In Experiment 1, the accuracy of recognizing characters with more nodes was higher and the response time was lower. The repeated-measures ANOVA of accuracy and RT found that the main effect of the number of nodes was significant. The interaction between the number of nodes and presentation time was significant. When the stimulus presentation time were 40 ms and 50 ms, node number effect was more pronounced. In Experiment 2, the accuracy of recognizing characters covering the high complex nodes was lower. The repeated-measures ANOVA of accuracy found that the main effect of the complexity of nodes was significant. The interaction between node complexity and node order is significant. Node complexity effect was more pronounced when covering the fifth nodes. In Experiment 3, we also found that the main effect of the complexity of nodes was significant. The interaction between node complexity and node generation method is significant.

These findings support the nodes provide bottom-up stroke separation guidance information. Stroke separation was performed in parallel, for the more nodes a character has, the more information provided for the stroke segmentation, and therefore the character would be easier to recognize. And stroke separating began with the extraction and analysis of nodes, for the more complex nodes are, the greater impact on recognize. This study deepens the understanding of the early visual process of Chinese character recognition and supports Chinese character recognition is a generative reverse reasoning process, which could contribute to develop a complete cognitive model of Chinese character recognition.

Keywords handwritten Chinese character recognition, node, stroke, generative model