

· 综述 ·

# 英文手写体字符、字词识别研究的现状<sup>①②</sup>

宋合义 荆其诚

西安交通大学管理学院 中国科学院心理研究所  
(西安 710049) (北京 100012)

[摘要]对人类识别英文手写体字符、字词的研究目前主要有两条途径。一是研究手写体字符、字词的静态特征对识别的影响;二是研究手写体字符、字词的动态特征对识别的影响。对静态特征的研究已表明,确定人类识别易混淆字符的物理—功能规则有助于提高机器识别英文字符、字词手写体的识别率。对动态特征的研究已表明,人类可以从英文手写体字符、字词的静态轨迹中知觉到动态信息,英文手写体字符、字词动态特征的变化对识别有显著的影响。

关键词 人类识别,英文手写体字符,静态特征,动态特征

了解人类识别手写体字符的认知规律是信息处理科学文字处理自动化领域给认知心理学研究提出的重要课题。这方面的研究不仅在实践上对手写体文字信息处理自动化的实现有所帮助,而且在理论上对了解人类的认知规律也是有意义的。为了适应计算机对手写体字符、字词识别的需要,也为了了解人类的认知规律,西方心理学家用英文字符手写体为材料对人类识别手写体字符进行了一些研究。这些研究目前来看主要有两条途径。一条是对手写体字符、字词的静态特征对识别的影响的分析;另一条是对手写体字符、字词动态特征对识别的影响的研究。

## 1 英文手写体字符、字词静态特征对识别的影响的研究

研究手写体字符静态特征对识别的影响的一个方面是分析手写体字符静态特征的变异因素。

Ward 和 Kuklinski 发现<sup>[1]</sup>,对于英文手写体而言,这些变异因素包括笔画连接、弯曲、圈和歧点(cusping)、起始线与折返线的消失、勾与连续、介于折返线与勾之间的形状、断笔。

研究手写体字符静态特征对识别的影响的另一个方面是人们在识别手写体字符时根据什么特征来区分容易混淆的字符。这方面研究的基本研究思路是:首先找出识别英文手写体字符时容易产生混淆的字符对,并建立混淆矩阵。在此基础上分析混淆字符对的混淆特征是什么,如 Y—V 的混淆特征是书写时右边斜线是否延伸及延伸的程度;U—V 的混淆特征是底部宽度的差异等。然后以这些混淆特征作为自变量研究人类对它们的区分阈限,再以这

① 本文于 1995 年 11 月 6 日收到。

② 国家自然科学基金资助项目。

个区分阈限作为机器识别混淆字符的区分界限。这方面的研究资料为提高计算机识别手写体字符的识别率做出了一定贡献。

Shillman, Kuklinski 和 Blesser 研究了人们如何在识别手写体字符时区分容易混淆的字符对 Y 与 V、F 与 C、H 与 U<sup>[2]</sup>, 并建立了研究手写体字符区分性特征的三种心理技术。他们认为, 手写体字母与标准的印刷体字母的形状有明显的不同。标准印刷体字母间有明显的边界, 而手写体字母间的边界从物理特性上看则不明显。他们研究的目的是想了解人类在识别手写体字母时如何确定在物理表现上容易混淆的字母的边界, 从而把容易混淆的字母区分开。

他们的研究基于他们提出的物理—功能规则(Physical to Functional Rule)理论。这种理论认为, 任何一个字符都具有物理特征、知觉特征和功能特征。字符的物理特征指该字符通常用几何或拓扑术语描述的那些部分, 它们组成物理图象, 如线、角、交叉。字符的知觉特征是指字符呈现时被知觉到的属性; 例如, 尽管一条线在物理上比另一条线长, 但这两条线有可能被知觉为相等。功能特征是提供了字母的一个基本描述的抽象。对手写体 V 与 Y 而言, 其物理特征的状态是通过对右边那条线的延伸出来的状态的物理测量决定的; 知觉特征的状态是由被试在实验中对“你是否看到右边那条线延伸出来了”的问题的回答而决定的; 功能特征状态只能由对“它是 Y, 还是 V”的问题的回答而决定。

实质上, 所谓的物理—功能规则就是确定易混淆的字母的边界, 具体地说, 是确定对具有一定物理形状的易混淆的手写体字符来说, 把它们识别为什么字母。

他们用下列三种实验技术研究了易混淆字母的物理—功能规则。

标定实验(Labeling experiments)。在这个实验中他们从手写印刷体的 V—Y 系列中选出了七个字符, 这七个字符的  $l_1$  与 L 的比率从 0 到 0.5 (见图 1), 制成七张刺激卡片。然后让被试识别这七张卡片, 哪个是 V, 哪个是 Y。然后分别计算被试对每张卡片回答为 V 和 Y 的百分比  $P_v(i)$  与  $P_y(i)$ 。对  $P_v(i)$  和  $P_y(i)$  作图。  $P_v(i)$  与  $P_y(i)$  两条曲线的交叉点即为 V 与 Y 两个字符的边界。研究结果表明 V 与 Y 的边界为  $l_1/L = 0.17$ , 则 V—Y 手写印刷体的物理功能规则系数为 0.17。即当  $l_1/L > 0.17$  时, 被试将其识别为 Y;  $l_1/L < 0.17$  时, 被试将其识别



图 1 V—Y 字符的  $l_1$  与 L 的比率

这三个研究者在标定实验中发现, 被试在识别某些刺激时所用的时间多于平均识别时间。通过数据分析发现, 对于交叉区域的刺激, 被试需要花更多的时间去标定。于是这三个研究者尝试用反应时实验的方法来区分字母边界。

反应时实验(Reaction time experiments)。在这个实验中, 研究者用单通道速示仪呈现前一个实验的七张卡片, 然后让被试读出每张卡片的字母名称, 记录从刺激呈现到被试读出字母名称的时间。结果发现, 在  $l_1/L = 0.17$  处的刺激, 被试的平均反应时显著增加, 这一结果与前一个实验的结果一致。因而研究者认为, 反应时是确定字母间边界的另一可行技术。

良好实验(Goodness experiments)。在这个实验中,研究者让被试用0—5级六点量表对七张卡片上的字符象V和象Y的程度分别进行主观评价。然后计算所有被试的平均良好评价等级 $G_v$ 和 $G_y$ ,并对 $G_v$ 与 $G_y$ 作图,其交叉点即为V与Y两个字母的边界,其求得的 $I_1/L$ 值基本上与前两个实验相同。

Shillmant等人的研究还表明<sup>[2]</sup>:V—Y手写体的物理—功能规则对于一个较大范围的角度(42度—90度)是恒定的。而且对C—F、U—H这两对易混淆的手写体字符进行的良好性实验也得到了各自的物理—功能规则。这些研究证明,在中性图形背景中,手写体字符的物理—功能规则是独立于交叉角、字母大小和字母标名(Label)的一个常数。

Suen用标定法研究了人对U、V这两个最易混淆的字母的不受限制的手写体的识别<sup>[3]</sup>,找出了人在识别这两个字母手写体时所依据的主要特征,并把这些特征编进计算机识别的算法中,使计算机对这两个字母手写体的识别率大大提高,达到94%。这一结果证明,基于心理学的特征与决策标准对于字母计算机识别的作用是很大的。

上述通过人类识别手写体字母的研究建立的物理—功能规则理论也为机器识别手写体字符在方法上提供了帮助。它提出的对易混淆的字符的识别方法导致了机器识别的健壮性识别方法<sup>[4]</sup>的产生。所谓健壮性是指一个系统,一旦发生故障就能立即恢复,并且它的反应速度足以使其很快摆脱随机错误。也就是说,软件或硬件中错误的严重性如何,主要取决于这个系统的自动排除错误和不给用户什么麻烦就能恢复工作的能力的一种设计<sup>[5]</sup>。

## 2 英文手写体字符、字词的动态特征对识别的影响的研究

所谓动态特征主要包括书写时的起落笔数、笔顺、每笔的书写方向以及每笔的书写速度。有关动态特征对英文手写体字符识别的影响的研究,目前主要涉及以下两个方面的问题:第一,人们是否能从已写好的手写体字符的静态轨迹中知觉到动态信息。第二,这种动态信息是否会影响人们对手写体字符的识别。

Babcock和Freyd的研究对第一个问题做出了回答<sup>[6]</sup>。在他们进行的隐含识别作业与明确识别作业的两个实验中,他们发现,被试可以从手写体的静态轨迹中抽取有关书写过程的动态信息。

Zimmer进行的研究也表明<sup>[7]</sup>,动态信息可能与手写体的识别有关。Zimmer要求被试在说出不同类型的字母的心理表象的同时想象字母的特点。他发现被试在说出一个正在书写的字母的表象时,要比他在说出一个已经写好的字母的静态表象时,更能精确地描述字母书写过程。作者认为,这意味着手写体的最便利的表征是包含有字母书写方法知识的表征。

Freyd是目前在动态特征对手写体字符识别的影响这一领域研究最多的学者。她在1983年开始研究,在实验中<sup>[8]</sup>,让被试看一些新奇的字符。这些字符是用两种书写方法(如,两种笔画方向)中的一种写在计算机屏幕上。随后要求被试辨认他们学过的字符的变形版和无变形版。结果表明,被试对那些没有变形的字符,以及那些由与他们学过的书写方法一致的书写方法产生变形的字符的识别成绩较好。另外,结果表明,若书写变形字符的方法是被试学过的方法,这种方法产生的变形字符的潦草线条是较好识别的。这意味着一些变形可能包含有用的信息。有可能是被试在学习作业期间所用的字符书写方法的知识帮助他们了对变形字符的识别。这些结果表明,当读者与书写者具有共同的书写系统时,他们的识别成绩

更好。据此 Freyd 提出了“共同方法”理论来解释动态特征对手写体字符识别的影响。这一理论认为,手写体的识别要使用有关字母是如何书写的信息。具体地讲,知觉者用他们的书写方法的知识,来从一个特定的手写体字母的静态轨迹中推论一个特定字母的书写过程的动态信息。如果读者具有原来写这些字符所使用的方法的知识,那么这些字符将是容易辨认的。更明确地说,这意味着学了同样一套字符书写方法的人之间比学习不同书写方法的人之间的文字交流障碍更少。

共同方法理论提出后,Freyd 等人进行了一些研究来验证她的理论。Babcock 和 Freyd 的研究表明<sup>[8]</sup>,识读者对表现在字符静态轨迹中的书写方法的信息是敏感的。Dekay 和 Freyd 的研究结果也表明<sup>[9]</sup>,被试对用他们自己的书写规则写出的字符对比用其它书写规则写出的字符对能够更好地进行区分。这些结果都为共同方法理论提供了证据。

但也有研究对共同方法理论进行了挑战,Suen 进行的研究就是一个例子<sup>[10]</sup>。在研究中,Suen 用让一组左利手和右利手的被试用三种字体[即印刷体(Printing)、手写体(Manuscript)和草体(Cursive)]写一个句子的方法收集了许多字符。这些句子被分割成单独的字母后,要求第二组左利手和右利手的被试辨认每个字母。如果可以认为左利手被试和右利手被试是用不同的方法书写字母的,那么根据共同方法假设,左利手被试对于左利手的书写者写出的字母识别得更好,而右利手被试对右利手的书写者写出的字母识别的更好。而这种左、右利手的书写者与左、右利手的识读者的交互作用没有被发现。但持共同方法假设观点的研究者认为,没有发现这种交互作用的原因可能是对左、右利手的书写者原本是否使用的是两种完全不同的方法并不清楚。

总之,有关共同方法假设的研究还没有得到一致的结果。看来,对共同方法假设的一个更强有力的验证应该用其书写方法已被了解的书写者与识读者,并使用一套已知是用两种不同的方法书写的一套字符。

对手写体字符识别的研究使一些研究者发现,在字母识别中占支配地位的特征分析理论不适合于手写体字母的识别,因为它们不能很有效地总结手写体字母的特征。在手写体字母中的特征常常是歪曲的,或者完全找不出原来的特征。由于特征分析理论不能对手写体字母的识别作出很好的解释,使一些研究者提出了分离加工系统理论。这一理论认为有两个独立的加工系统分别负责印刷体和手写体的识别。

Corcoran 和 Rouse 提出了证据来支持这种分离加工系统的观点<sup>[11]</sup>。他们用速示仪测量了以分组方式呈现和以混合方式呈现的字词的识别正确率。在这个实验中,他们进行了三种刺激的比较:印刷体字词与手写体字词的比较;两个不同的书写者写出的两种字体的比较;印刷体大字母的字词与印刷体小写字母的字词的比较。研究结果表明,在混合呈现方式下,只有印刷体与手写体的比较产生了衰减。这一结果反映了被试需要为每一种呈现建立适当的策略。因此,这一结果支持分离加工系统理论。

Babcock 和 Freyd 也认为<sup>[8]</sup>,在字母识别中占支配地位的特征分析理论不适合于手写体字母的识别。于是他们在动态特征对手写体识别有显著的作用的实验结果的基础上,对特征分析理论做了修正,提出了互补识别机制的理论。这一理论认为,一个利用手写体静态轨迹中的内在书写过程的信息的加工,可能与分析结构信息的更传统的特征分析加工过程共同发生作用。也可能是前一个机制主要用于阅读手写体字母,而后一个机制主要用于阅读印刷

体字母。每一个机制的使用范围可能依赖于字母静态结构的一致性或依赖于代表这些字母所用的书写方法的变形的一致性。

对手写体字符、字词识别的动态特征的研究为机器识别手写体字符、字词提供了帮助。它导致机器识别模型的产生。这一模型是通过把笔尖的运动描述成一个时间函数的数学模型对笔划进行描述。这样,一个手写体字词就可以被分成笔划,并用模型参数和字母序列对笔划进行分类,在这个基础上对字词进行识别。

值得一提的是,一些机器视觉字符识别程序已经使用了轮廓和主干追踪模型。这些模型试图通过连续追踪原来的书写过程来确认字符<sup>[12]</sup>。这种模型可以利用驻存在字符静态轨迹中的动态信息。

从事计算机手写体识别的研究工作者 Hoffman, Skrzypek 和 Vidal 也认为手写体字符的识别加工应该探索字符书写机制的知识<sup>[13]</sup>。手写字母是由有限数量的手的运动形成的。这些动态信息被人们贮存在记忆中,并且既能在认知上被用来对在视觉上被感觉到的字符进行分类,也能在认知上被用来在书写时控制运动器官的运动。他们在机器上进行的实验表明,由字符书写机制所构成的识别手写体的方法是有用的。它可能比那些把字符切分成单个笔画或任何较低水平的视觉特征的识别方法更有希望,因为任何一套给定的基本笔画的完整性往往不确定。

## 参考文献

- [1] Ward J, Kuklinski T T. A model for variability effects in handprinting with implications for the design of handwriting character recognition systems. *IEEE Transaction on System, Man, and Cybernetics*. 1988, 18, 438—451.
- [2] Shillman R J, Kuklinski T T, Bleaser B A. Psychophysical techniques for investigating the distinctive features of letters. *International Journal of Man—Machine Studies*, 1976, 8; 195—205.
- [3] Suen C Y, Shillman R J. Low error rate optical character of unconstrained handprinted letters based on a model of human perception. *IEEE Transactions on Systems, Man and Cybernetics*. 1977, SMC-7; 491—495.
- [4] Tappert C C, Suen C Y, Wakahara T. The state of the art in on—line handwriting recognition. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 1990, 12; 787—808.
- [5] 夏培肃主编. 英汉计算机辞典. 北京: 人民邮电出版社, 1984.
- [6] Babcock M K, Freyd J J. Perception of dynamic information in static handwritten form. *American Journal of Psychology*, 1988, 101; 111—130.
- [7] Zimmer A. Do we see what makes our script characteristic or do we only feel it? *Psychological Research*, 1982, 44; 165—174.
- [8] Freyd J J. Representing the dynamics of a static form. *Memory & Cognition*, 1983, 11; 342—346.
- [9] Dekay M L, Freyd J J. The effects of drawing method on the discriminability of characters. *Visible Language*, 1991, 25; 377—414.
- [10] Suen C Y. Handwriting generation, perception, and recognition. *Acta Psychology*, 1983, 54; 295—312.
- [11] Corcoran O W J, Rouse R O. An aspect of perceptual organization involved in reading typed and handwritten words. *Quarterly Journal of Experimental Psychology*, 1970, 22; 526—530.
- [12] Ullmann J R. A use of continuity in character recognition. *IEEE Transaction on System, Man and Cybernetics*, 1974, 4; 294—300.
- [13] Hoffman J, Skrzypek J, Vidal J J. Cluster network for recognition of handwritten, cursive script characters. *Neural Networks*, 1993, 6; 69—78.