

# 基于空间分割的手写输入系统的用户绩效模型<sup>\*</sup>

吴昌旭<sup>1</sup> 李怀龙<sup>2</sup> 杨群会<sup>2</sup> 张侃<sup>2</sup>

(<sup>1</sup>University of Michigan, Ann Arbor, U. S. A. 48109) (<sup>2</sup>中国科学院心理研究所, 北京, 100101)

**摘要** 为定量估计与提高基于空间分割的手写输入系统用户绩效, 运用 Fitts 定律及离散变量的数学期望等方法推导了用户绩效的数学模型。实验验证了该模型在具体界面中能较好地拟合用户在 8 次(8 × 24 字)训练后的实际操作绩效。应用该模型可较好地拟合用户使用手写系统完成抄写任务的时间、推算输入系统某些参数的最佳设置值; 并发现当输入系统的其它参数为定值时, 缩短用户单字手写时间会比缩短系统识别时间能更有效地提高用户绩效。

**关键词** 数学心理模型, 手写汉字输入系统, 绩效模型, 空间分割。

**分类号** B849:TB18

## 1 引言

随着便携式电子设备发展, 笔计算与笔交互在便携式电子设备中可能取代传统键盘, 成为当今世界人机交互的新的手段与研究热点<sup>[1~4]</sup>。以往汉字输入的工程心理学研究主要集中在键盘输入<sup>[4]</sup>, 而对汉字手写输入系统的工程心理学研究仍属空白。手写输入系统的用户绩效是决定用户对系统满意感与可接受程度的关键<sup>[5,6]</sup>。建立用户绩效模型、弄清用户完成任务的整个过程由哪些子过程构成、子过程中的各因素与绩效的定量关系是什么, 是分析与提高这一绩效的基础。通过建立绩效模型可以估计用户在特定手写输入系统下完成任务的时间, 为系统设计提供定量的设计建议。

目前 PDA 及其他手写系统的界面大多采用基于空间分割的手写输入方法, 即利用字与字之间的空间距离分割连续手写的汉字(见图 1)。那么能否对这些基于空间分割的手写系统进行用户绩效的建模呢? 此外, 手写系统设计人员正投入大量的人力、物力缩短系统的识别时间, 那么是否能通过模型发现比识别时间更为重要的影响用户绩效的因素? 改善这些因素可能会更有效地提高用户绩效。

理论上系统绩效往往包括任务完成时间与错误率。Casali 和 Williges(1990) 对语音输入系统的研究将任务完成时间与未要求修改的错误量作为绩效的

指标<sup>[7]</sup>。在本研究中, 因为被试必须修改所有错误识别的汉字, 错误识别的字数越多, 被试完成整个任务的时间就越长, 所以任务完成时间本身已包含了错误量这一因素, 因而在本研究中仅将任务完成时间(D)作为系统绩效的指标。

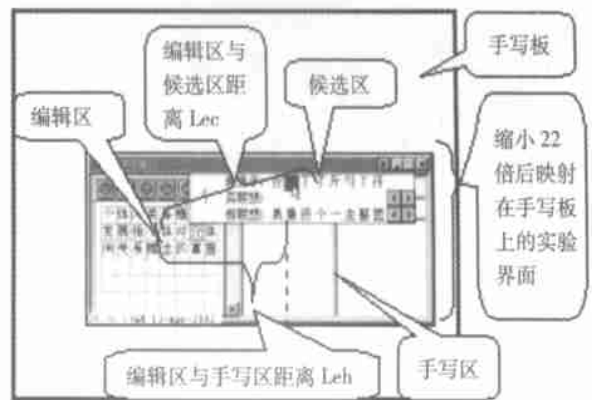


图 1 基于空间分割的手写输入系统用户界面  
(屏幕上的界面尺寸按 2.2 倍增益缩小到手写板上) 示意图

## 2 用户绩效模型的推导

用户与这种基于空间分割的手写输入系统的交互包括连续手写与再修改两个过程。因此, 用户的绩效模型也由这两部分组成: 手写输入部分与修改部分。

### 2.1 输入部分用户绩效模型的推导

手写输入的过程, 包括书写笔在输入框之间移

收稿日期: 2003 - 04 - 16

<sup>\*</sup> 本研究获得英特尔中国研究院(Intel China Research Center)的资助。

通讯作者: 张侃, E-mail: zhangk@psych.ac.cn

动、手写汉字、用户读取下一个待输入汉字三个子过程。

**2.1.1 书写笔在输入框之间的移动时间(MT)** 根据 MacKenzie (1992) 对 Fitts Law (手指或手持笔在两个目标间移动的时间与移动距离与目标宽度的函数关系) 在人机交互中应用的总结<sup>[8]</sup>, 手持指点工具在两个目标间移动的时间与使用的定位工具及身体部位有关, 不同的定位工具 (触笔、游戏杆、鼠标等) 与身体部位有不同的 Fitts 方程参数。从手写笔的操作与定位方式看, 本实验中使用的手写触笔与 Fitts (1954) 使用的定位工具触笔极为相似, 两者的手持方式、人手的运动控制部位相同, 而且定位装置均属于 0 阶位置控制<sup>[9,10]</sup>; 主要不同之处在于: 1) 在 Fitts (1954) 的实验中, 被试使用触笔直接在两块导电板间敲击, 即运动的视觉反馈与运动操作在同一界面上 (On - display); 而本研究中视觉反馈是在屏幕上, 而运动操作在手写板上完成, 两者在不同的界面上 (Off - display)<sup>[11]</sup>; 2) 在 Fitts (1954) 的实验中, 触笔的运动操作轨迹与触笔视觉运动轨迹是点对点对应的, 操作与视反馈的幅度增益值 (Gain) = 1; 而本研究中手写触笔的运动操作幅度与其视觉运动幅度成一定的比率关系: 手写触笔在手写板上每移动 1cm, 屏幕上对应的光标移动 2.2 cm, 即操作与视反馈的幅度增益值为 2.2。

但是, 已有研究表明: (1) On - display 与 Off - display 的手动 0 阶位置控制设备在用户绩效上无显著差异<sup>[11]</sup>; (2) 大量研究表明操作与视反馈的幅度增益值发生较大变化时, 用户操作时间并未受到影响<sup>[1,12,13]</sup>; 所以, 用户使用本研究中的手写实验装置进行目标定位时, 与 Fitts (1954) 的实验中得到的数据满足相似的绩效模型。因此, 本研究中直接采用 Fitts (1954) 的实验中得到的 Fitts 方程作为预测用户使用手写板进行目标定位时间的方程。这一方程为:

$MT = 12.8 + 94.7 \log_2(2A/W)$  (MT: 移动时间; A: 目标间距; W: 目标宽度)

Evan & MacKenzie (1995) 的研究表明, 对手动操作的绩效建模应以手部动作的运动幅度为建模对象, 而不是屏幕上的光标等视觉反馈的运动幅度<sup>[14]</sup>。所以, 将屏幕上的界面尺寸按 2.2 倍 (增益为 2.2) 缩小到手写板上 (见图 1), 从而测量实际手部在手写板上目标间移动的距离与目标在手写板上的宽度。

设在输入过程中, 共输入 N 个汉字。手写区输

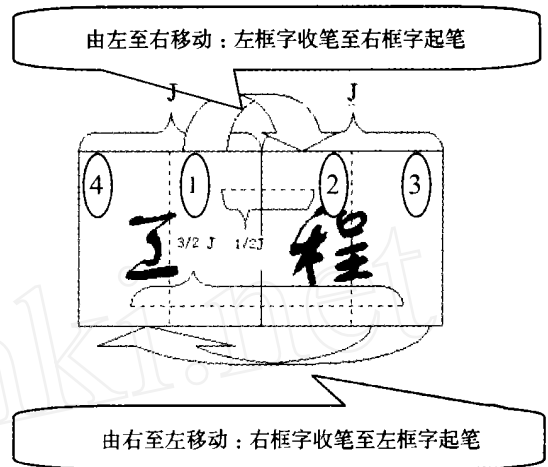


图 2 手写区示意图

入框 (正方形) 的内边长为 J。对于双框手写界面, 手写过程中, 书写笔在两框内来回连续移动输入汉字。由于汉字的书写顺序是由左至右, 大多数汉字都从左半区域起笔、右半区域收笔, 在手写识别系统的界面中, 书写笔在两框内来回连续移动时, 见图 2:

(1) 由左至右移动: 由图 2, 当左框字收笔至右框字起笔, 书写笔由左至右移动时 (区域 - > ), 平均目标间距为  $(1/2)J$ , 目标宽度为  $(1/2)J$ , 移动的时间 ( $MT_1$ ) 为:

$$MT_1 = 12.8 + 94.7 \log_2[2(1/2)J / (1/2)J] = 107.5 \text{ ms}$$

(2) 由右至左移动: 由图 2, 当右框字收笔至左框字起笔, 书写笔由右至左移动时 (区域 - > ), 平均目标间距为  $(3/2)J$ , 目标宽度为  $(1/2)J$ , 移动的时间 ( $MT_2$ ) 为:

$$MT_2 = 12.8 + 94.7 \log_2[2(3/2)J / (1/2)J] = 257.6 \text{ ms}$$

所以, 每字平均移动时间  $MT = (1/2)(MT_1 + MT_2) = 183 \text{ ms}$ ; 当推广到框数 = F 时, 书写 N 个汉字:

$$MT_1 = (F - 1) [12.8 + 94.7 \log_2(2(J/2) / (J/2))] = (F - 1)(12.8 + 94.7);$$

$$MT_2 = \{12.8 + 94.7 \log_2[2 \times (F - J/2) / (J/2)]\}$$

用户花费在框与框之间移动的时间:  $MT_F = (N/F) \times (MT_1 + MT_2) = (N/F) \times [107.5(F - 1) + 94.7 \log_2(4F - 2) + 12.8]$

**2.1.2 手写汉字时间(WT)** 由平均单字手写时间 (WT) 确定。

**2.1.3 用户读取下一个待输入汉字时间(VD)** 根

据张侃(1991)提出的汉字键盘输入认知模型,汉字输入无论是听打、看打、还是想打,都经历知觉加工(视、听觉存储转换为音、形表象)、认知加工(对汉字进行空间、音素与键位分析)与动作加工三个加工过程<sup>[15]</sup>。而手写输入用户,无论是听写、看写(本研究实验任务)还是想写,读取下一个待输入汉字也经历了上述过程中的知觉加工过程;反映这一知觉加工时间的研究是高定国(1995)对汉字的认知速度研究,他发现高频少笔划汉字单字判读(见字读音)反应时均值为439ms,标准差为48ms<sup>[16]</sup>;因为本研究实验材料也为高频少笔划汉字,所以将这一均值作为用户读取下一个待输入汉字的估计时间。综合三个子过程, $D_{input} = N \times (WT + MT + VD)$ ;双框输入界面: $D_{input} = N \times (WT + 183 + 439)$ ;推广到框数 = F 时, $D_{input} = N \times \{WT + 439 + (N/F) \times (MT_1 + MT_2)\} = (N/F) \times [107.5(F - 1) + 94.7 \log_2(4F - 2) + 12.8]$

(1)

## 2.2 修改部分的绩效模型的推导

连续手写输入结束后,用户对错误识别汉字的修改依据不同的识别结果将经历以下5条路线:

(路线A)在手写输入结束后,用户首先在编辑区寻找并用手写笔定位出错的汉字(见图1);然后,用户在候选区寻找正确候选字是否已出现在候选区中;确认出现后,手写笔将从编辑区移到候选区,在候选区中定位正确候选字;最后,手写笔将从候选区返回到编辑区进行下一个汉字的修改操作。设书写N个汉字的过程中有n<sub>1</sub>个汉字的修改经历了路线A。

(路线B)定位错误识别的汉字后,如果未在候选区找到正确候选字后(见图1),手写笔将从编辑区移到书写区,在书写区重写该字(第一次);待该字被识别后,手写笔将从书写区返回到编辑区。设书写N个汉字的过程中有n<sub>2</sub>个汉字的修改经历了路线B。路线B之后又将分为B<sub>1</sub>,B<sub>2</sub>,B<sub>3</sub>三条子路线:(子路线B<sub>1</sub>)第一次重写汉字被识别后,用户必须对这一单字识别结果进行判断,如果结果未被正确识别,则在候选区寻找正确候选字;如果还是未能在候选字中找到正确候选字,手写笔将从编辑区移到书写区重写该字(第二次)(包括识别时间R);用户判断第二次手写结果正确与否,如果正确则修改结束;如果仍错误,则根据指导语修改过程也将终止;设书写N个汉字的过程中有n<sub>3</sub>个汉字的修改经历了路线B<sub>1</sub>。(子路线B<sub>2</sub>)如果用户在候选区中找

到了正确候选字,手写笔将从编辑区移到候选区,在候选区中定位正确候选字;最后,手写笔将从候选区返回到编辑区进行下一个汉字的修改操作。设书写N个汉字的过程中有n<sub>4</sub>个汉字的修改经历了路线B<sub>2</sub>。(子路线B<sub>3</sub>)如果第一次重写的结果被正确识别了,用户判断结果正确即可。书写N个汉字的过程中将有n<sub>2</sub> - (n<sub>3</sub> + n<sub>4</sub>)个汉字的修改经历了路线B<sub>3</sub>。

由于篇幅限制,这里以路线A为例,推导对路线A的修改时间T<sub>A</sub>。其余路线的推导结果见附录3。

**2.2.1 路线A分析** 路线A包括5个子过程与时间:

(1)在编辑区定位错误识别汉字的时间(Def/ Duration of pinpointing in Edit Field)

包括在编辑区寻找到错误识别的汉字时间(D<sub>vef</sub>)与用手写笔选中该汉字的时间(D<sub>mef</sub>)。

由于是在输入结束后立即进行文本修改,而且在输入过程中,联机手写系统会实时呈现给用户识别结果,因此用户对编辑区的文本已发生的错误识别有一定的熟悉与记忆。在这种情况下,在N个汉字中寻找到一个汉字的时间估计为:

对N个简单刺激的反应时RT = 172.766 + 199.037ln(N) (推导见附录1)

(2)

加上对高频少笔划汉字的判读时间(439ms)。即, $D_{vef} = (173 + 199 \ln(N)) + 439 = 612 + 199 \ln(N)$

(3)

采用离散变量求数学期望的方法<sup>[19]</sup>,在已知编辑区尺寸(a × b个单位)的情况下,可以推算出手写笔选中随机出现在编辑区的错误识别汉字的平均移动距离,再用Fitts Law估计选中时间,得:

$D_{mef} = 12.8 + 94.7 \log_2 \{2 \times 1 / (ab) \times [(0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2]^{0.5}\}$  (推导见附录2)

(4)

在编辑区定位错误识别汉字的时间Def = D<sub>vef</sub> + D<sub>mef</sub> = 612 + 199ln(N) + D<sub>mef</sub> = 612 + 199ln(N) + 12.8 + 94.7log<sub>2</sub>{2 × 1 / (ab) × [(0.5a - m + 0.5)<sup>2</sup> + (0.5b - n + 0.5)<sup>2</sup>]<sup>0.5</sup>}

(5)

当a = 8, b = 3, N = 24时, D<sub>mef</sub> = 216ms, Def = D<sub>vef</sub> + D<sub>mef</sub> = 612 + 199ln(24) + 216 = 1460ms

(6)

(2)在候选区寻找正确候选字的时间(D<sub>vcf</sub>/ Duration of Visual searching in Choice Field)

设候选字数为I,仿上,在I个汉字中寻找到一个

个汉字的时间刻估计为:对 I 个简单刺激的反应时加上对高频汉字的认知时间(439ms)。即,  $D_{vcf} = (173 + 199\text{Ln}(I)) + 439 = 612 + 199\text{Ln}(I)$  (7)

(3) 从编辑区到候选区的移动时间(Dec/ Duration of movement from edit field to choice field)

设编辑区到候选区的距离为 Lec (Length from edit field to choice field); 编辑区与候选区的平均宽度为 E(Wec), 则由 Fitts Law,  $\text{Dec} = 12.8 + 94.7\log_2(2A/W) = 12.8 + 94.7\log_2(2\text{Lec}/E(\text{Wec}))$  (8)

本研究经测量手写板上依据增益放大的用户界面尺寸,得:

$\text{Lec} = 29\text{mm}$ ,  $E(\text{Wec}) = 20\text{mm} = > \text{Dec} = 12.8 + 94.7\log_2(2 \times 29/20) = 158\text{ms}$

(4) 在候选区定位正确候选字的时间( $D_{mcf}$ / Duration of movement in choice field)

设长方形编辑区( $c \times d$ )中每个汉字的尺寸为 1 个单位,根据(3)、(4)式,得:

$D_{mcf} = 12.8 + 94.7\log_2\{2 \times 1/(ab) \times [(0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2]^{0.5}\}$  (9)

本研究中  $c = 8$ ,  $d = 1$  为例,  $E(L_{mcf}) = 2$ ; 在候选区用手写笔选中该汉字的时间  $D_{mcf} = 202\text{ms}$  (10)

(5) 从候选区返回到编辑区的移动时间(Dec/ Duration of movement from choice field to edit field)

由(3)  $\text{Dec} = 12.8 + 94.7\log_2(2\text{Lec}/E(\text{Wec})) = 158\text{ms}$  (11)

综合上述 5 个子过程,路线 A 总体耗时  $T_A$ :

$T_A$  一般式:  $T_A = D_{ef} + D_{vcf} + \text{Dec} + D_{mcf} + D_{ce} = D_{ef} + D_{vcf} + 2 \times \text{Dec} + D_{mcf} = 612 + 199\text{Ln}(N) + 12.8 + 94.7\log_2\{2 \times 1/(ab) \times [(0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2]^{0.5}\} + 612 + 199\text{Ln}(I) + 2 \times 12.8 + 2 \times 94.7\log_2(2\text{Lec}/E(\text{Wec})) + 12.8 + 94.7\log_2\{2 \times 1/(ab) \times [(0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2]^{0.5}\}$  (12)

加入本研究实际的界面参数后:  $T_A = D_{ef} + D_{vcf} + \text{Dec} + D_{mcf} + D_{ce} = 3002\text{ms}$

**2.2.2 综合修改与手写输入过程与时间** 类似路线 A,运用 Fitts 定律等得出路线 B,  $B_1, B_2, B_3$  的修改时间的一般式及它们在本研究中的具体估计值(见附录 3),乘以它们各自出现的频次,得到修改时间 T 的一般式与总时间的估计值(见附录 4)。

加入实际界面参数,双框手写界面输入耗时

$D_{input} = N \times (\text{WT} + 183 + 439)$ , 及  $N = 24$ ,  $I = 8$ ,  $R = 550\text{ms}$  后:

$D = 3002 \times n_1 + (3339 + \text{WT}) \times n_2 + (2757 + \text{WT}) \times n_3 + 1983 \times n_4 + 439 \times (n_2 - (n_3 + n_4)) + 24 \times (\text{WT} + 439 + 183)$  (15)

### 3 用户绩效模型的验证实验

为验证该绩效模型(15)对实际数据的拟合能力,在手写输入系统的新手在学习使用手写系统的过程中,考察用户完成任务的时间,并同时纪录错误发生频次( $n_1, n_2, n_3, n_4$ )、测试被试的手写速度(WT)。将这些参数代入绩效模型(15)中,求出绩效模型对用户完成任务时间的理论估计值。之后使用决定系数等来考察绩效模型的估计值对完成任务的时间的解释率。

#### 3.1 实验方法

**3.1.1 实验材料** 在心理学百科全书上随机抽取心理统计学的术语 18 条,略加修改后保证字数相同(每句 24 个汉字),总笔划数基本一致(均值 179 笔/句,标准差 11 笔/句),均为高频少笔划汉字。通过预实验挑选出 16 句,在将手写速度作为协变量时,10 名非心理或统计专业的三年级大学生对这 16 句话的抄写时间无显著差异( $F(15, 60) = 0.567$ ,  $p = 0.888 > 0.05$ )。

**3.1.2 实验设备** (1) 手写硬件设备:汉王 99 手写输入系统(手写板、识别系统)一套,型号 EM III 6045。(2) 实验手写软件设置识别时间(R) 550ms;(3) 手写速度测试软件与方法:采用自带手写速度纪录的 IntelPen 汉字识别软件。计时精度高于 50ms。因为该软件对手写字大小不作限制,所以为了避免手写字的大小对手写速度测试的影响,在手写速度测试时,将印有汉王手写界面的透明胶片粘贴在屏幕上,要求被试仅在胶片所印的界面中的左框中进行手写输入。(4) 数据纪录与分析软件:Snagit/ 32 (4.3.6 版) 实时屏幕摄像软件一套,实验时以 5Hz/second 的频率采集屏幕手写操作情况,生成的 AVI 文件存储于主机硬盘。Premiere AVI 视频分析软件一套,对 Snagit 生成的 AVI 文件进行后期处理。

**3.1.3 实验设计与实验过程** 实验采用被试内单因素设计。自变量为被试的手写训练次数,其余变量为软件纪录。实验材料的呈现采用 ABBA 顺序。待抄写文本为 8 行  $\times$  3 列呈现,与识别软件编辑区 8 行  $\times$  3 列对应。

首先向被试介绍试验用手写软件的界面与使用方法,包括连续的手写输入、如何选中编辑区文本进行修改,如何在候选区选择被选汉字等。并给出指导语:“1)请您按平时的书写习惯以最快速度抄写屏幕上的话;2)请您先抄写完整个句子,然后再修改错误识别的字;3)修改时请先看一下候选字中是否有正确候选汉字,如果没有,那么需要您重写该字。如果一个字重写二遍还是无法识别,请您跳过。”

然后每个被试采用各自用户名登录识别软件,熟悉手写板定位、输入、修改的基本操作,包括握笔姿势、点击动作、手写笔在书写板的位置相对与光标在屏幕上的位置、尝试连续在双框进行手写输入、在编辑区与候选区选中目标汉字等。之后,被试进入正式16次训练(预实验表明,16次 $\times$ 24字/次=384字的训练可以使用户完成24字输入与修改任务的时间 $<100$ s)。训练分为4个阶段,每个阶段各4次训练。在每次训练之后,被试有15s左右休息时间。整个训练在1小时内完成。最后,休息3min后,使用Intelpen软件及相应测试方法(见实验设备)对用户的手写速度进行测试。

**3.1.4 被试** 大学二年级学生6人(男女各半),年龄20~22岁。所有被试矫正视力正常,实验过程中均采用右利手进行实验。所有被试以前未使用过任何手写输入设备与系统,来自非统计与心理专业,不熟悉统计学知识。

### 3.2 实验结果

**3.2.1 4个训练阶段实际任务完成时间与模型估计值的相关分析** 采用贡献率/决定系数作为考察模型对实际数据拟合指标<sup>[20,21]</sup>。

即随着训练次数的增加,模型估计值与实际任务完成时间的相关系数增大(见表1),模型估计值对实际任务完成时间的解释率(决定系数)也不断上升。

表1 4个训练阶段实际任务完成时间与模型估计值的相关系数与决定系数

训练阶段	实际任务完成时间与模型估计值的 Pearson 相关系数 $R$	$R^2$
1	0.678 **	0.460
2	0.741 **	0.549
3	0.925 **	0.856
4	0.934 **	0.872

注: \*\*  $p < 0.01$  (2-tailed)

**3.2.2 4个训练阶段实际任务完成时间与模型估计值的差异检验** 这4个训练阶段实际任务完成时间与模型估计值的比较见图3。

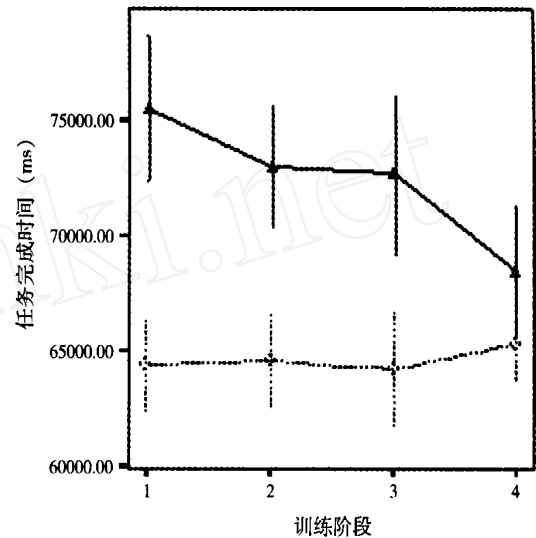


图3 4个训练阶段模型估计时间(虚线)与实际任务完成时间的差异(实线)比较(平均值与 $\pm$ 一个标准误差)

(1)阶段1(第1~4次训练):因为模型估计值与实际任务完成时间在这一阶段内为符合正态分布(Kolmogorov - Smirnov 检验  $p > 0.05$ ),故采用两独立样本的  $t$  test,检验实际任务完成时间与模型估计值的差异: $t(df = 44) = 2.559, p < 0.05$ ;说明前4次训练时,模型估计值尚与实际值有显著差异;(2)阶段2(第5~8次训练):因为模型估计值与实际任务完成时间在这一阶段内为不符合正态分布(Kolmogorov - Smirnov 检验  $p < 0.05$ ),故采用两独立样本的非参数检验考察实际任务完成时间与模型估计值的差异:Mann - Whitney  $U = 126$ ; Wilcoxon  $W = 379$ ;  $Z = -2.723$ ;  $p < 0.01$ ;说明前5~8次训练时,模型估计值仍与实际值有显著差异;(3)阶段3(第9~12次训练):因为模型估计值与实际任务完成时间在这一阶段内为符合正态分布(Kolmogorov - Smirnov 检验  $p > 0.05$ ),故采用两独立样本的  $t$  test,检验实际任务完成时间与模型估计值的差异: $t(df = 39) = 1.875, p > 0.05$ ;说明9~12次训练时,模型估计值已与实际值无显著差异;(4)阶段4(第13~16次训练):因为模型估计值与实际任务完成时间在这一阶段内为符合正态分布(Kolmogorov - Smirnov 检验  $p > 0.05$ ),故采用两独立样本的  $t$  test 检验实际任务完成时间与模型估计值的差异: $t(df = 46) = 0.912, p > 0.05$ ;说明最后4次训练时,模型估计值也与实际值无显著差异。

因此,随着训练次数的增加,模型估计值与实际任务完成时间的差异减小,最后两个训练阶段模型估计值与实际任务完成时间的差异已不显著,且模型估计值对实际任务完成时间的解释率(决定系数)较高。

## 4 模型的应用

建立了绩效模型后,手写系统设计者只要将已知的参数代入一般式(1)、(14)式,即可估计用户在一定手写输入界面与系统参数条件下完成抄写任务的时间。在根据特定任务设定用户绩效的最小提高值(如 200ms/24 字)后,就可以由绩效模型推算出系统参数的最佳值。例如,在其它参数不变的情况下,经模型计算,当手写框数(F)为 3,4,5 个等不断增加时,用户每书写 24 个字耗时会分别比手写框数为 2 时节省:41ms(F=3); 204ms(F=4); 357ms(F=5); 488ms(F=6); 597ms(F=7)…。所以,与双框手写界面相比,所有 4 框以上(含 4 框)的手写用户界面都能有效地节省用户在手写时消耗在框与框之间移动的时间(节省时间大于 200ms/24 字)。因此,在屏幕空间允许的前提下,建议采用 4 框或更多输入框的设计为宜。3 框手写界面节省的时间很少(41ms < 200ms),又因为三框界面要比双框占用更多的屏幕空间。这就提示系统设计者,在屏幕空间仅能容纳 2~3 框时,建议采用双框设计。

进一步应用模型,在已知界面参数的情况下, $N=24$ ,该界面条件下输入 24 个汉字的总共耗时为:模型(15)(修改时间)加上(1)(双框界面  $F=2$ ,输入时间),并化简得:

$$D = 199(n_1 + n_2 + n_3 + n_4) \ln(I) + (n_2 + n_3 + 24) WT + (n_2 + n_3) R + 2590n_1 + 2925n_2 + 2339n_3 + 1569n_4 + 439(n_2 - n_3 - n_4) + 24VD + 24 \times 183$$

因此,当识别错误的发生频率  $n_1, n_2, n_3, n_4$  (即系统的正确识别率)及候选汉字个数(I)为定值时,WT, R 与 D 呈线性增函数关系。又因为 WT, R 单位相同(均为 ms),所以它们各自的参数( $n_2 + n_3 + 24$ ), ( $n_2 + n_3$ )代表了这两个因素对因变量的权重。因此,两者对 D 的影响权重次序为:WT > R,即缩短用户的单字手写时间(WT)比缩短系统的识别时间(R)会对用户绩效产生更大的影响,并且随着输入字数 N 的增加,用户单字手写时间的权重还会逐渐增大。这就提示手写系统设计者,通过降低手写板与笔之间的摩擦系数等方法缩短用户的单字手写时间来提高用户绩效,会比通过优化算法等方法

缩短系统的识别时间更为有效。

## 5 讨 论

实验验证了基于空间分割的绩效模型能够较好地拟合用户在 8 次(24 × 8 字)训练后的实际操作绩效。其原因可能在于:由于模型在推导时多次使用 Fitts 定律对手写笔移动时间进行估计,而 Fitts (1954)的实验通过 613~2669 次移动的数据才获得 Fitts 方程的触笔表达式:  $MT = 12.8 + 94.7 \log_2(2A/W)$ <sup>[10]</sup>;也就是说,通过大量的移动训练才能达到与该表达式相似的移动时间。所以,只有当被试经历了前 8 次训练后,其用户绩效才能被 Fitts 定律准确估计。

但是,模型是假设用户按串行序列完成任务,而实际交互过程中有并行完成任务的可能。模型的建立采用类似 COMS 与 KLM 的人机交互分析方法。John (1996)认为 COMS 方法,及其各种变形(包括 NCOMSL, KLM 等)都很难考虑用户平行/同时进行两个及更多的任务<sup>[17]</sup>。例如,用户在进行路线 A 的子过程 2(在候选区寻找正确候选字)的同时,其手写笔已同时从编辑区向候选区的移动(路线 A 的子过程 3)。从这一点来看,模型可能高估了用户完成任务的时间。从实验的数据看,当训练次数在 14, 15, 16 次时,已有少量实际值与估计值的差值为负值。说明用户消耗了比模型估计更少的时间完成了任务。当然,有可能是用户在输入等操作上更为熟练,其平均动作时间小于模型使用的平均个体的参数,也可能正是由于上述平行加工的出现。

John(1989)<sup>[22]</sup>对打字员绩效模型的研究认为,手写过程中用户在读取屏幕信息的同时,也在进行动作输出。手写(边看边写)与打字(边看边打)相似,视觉信息读取与动作输出在日常工作中已获得大量的练习。因此,有可能视觉信息读取与动作输出的平行加工过程在整个实验阶段都是存在的。与之相反,用户在学习灵活地使用与控制输入笔时却需要一定时间的训练(因为输入笔与手写板之间的摩擦系数小于平时用笔与纸张之间的摩擦系数,而且输入笔的笔杆末端有一信号输出线)。在实验的前阶段(1~8 次训练左右),用户在学习灵活地使用与控制输入笔(训练因素),与此同时平行加工的过程也是存在的(平行加工因素)。但这时训练因素所引起的抄写时间的延长可能大于平行加工因素所带来的用户抄写时间的缩短。随着实验训练次数的增加(8 次训练后左右),训练因素对用户抄写时间

的影响逐渐减小,平行加工因素对抄写绩效的促进相对增加。也就是说,较为稳定的平行加工因素与逐渐变小的训练因素在实验过程中的叠加效应可能造成了实验整个过程中(前8次训练后与后8次训练)抄写绩效的变化。

为了解决对平行加工的建模问题,可尝试采用John提出的GOMS的另一变形方法CPCOMS对模型进行修改<sup>[17]</sup>;也可使用刘己力(1997)<sup>[23]</sup>提出的基于排队论的QN-MHP(Feyen, 2002)<sup>[24]</sup>进行建模:将实验参数作为随机变量并用系统仿真的方法建立Promodel计算机模型。

此外,本研究在考虑手写输入过程中未考虑到启动等汉字判读其他认知过程。因为手写内容存在着意义上的联系,因而可能存在启动效应。所以,在实验中的汉字判读的时间可能应为高频少笔划汉字认读时间(VD=439ms)减去上下文的启动效应所节省的时间,再加上判断与决策时间;但这两个子时间均需要进一步的实验研究。

## 6 总 结

(1) 本研究推导了基于空间分割手写输入系统用户绩效的数学模型(1)(输入部分)与(14)(修改部分)(见附录4),用户经过8次(8×24字)训练后,该模型能较好地拟合用户的抄写任务完成时间( $R^2 > 0.85$ )。

(2) 模型的应用:a)可估计用户在一定手写输入界面与系统参数条件下完成抄写任务的时间,并推算出系统某些参数的最佳值;b)发现当识别错误发生频率及候选汉字个数为定值时,缩短用户单字手写时间会比缩短系统识别时间能更有效地提高用户绩效。

## 参 考 文 献

- I. Scott M, Larry Ch. A Performance Comparison of Two Handwriting Recognizers. *Interacting with Computers*, 1999, 11(3): 283 ~ 297
- Richard C D, James L, Jason A B, et al. A Framework for Sharing Handwritten Notes Demonstrations. *Proceedings of the ACM Symposium on User Interface Software and Technology*. New York, USA: ACM Press, 1998. 119 ~ 120
- Heiko S, Ting T H, Gareth L. Beyond Translation: Approaches to Interactive Products for Chinese Consumers. *International Journal of Human - Computer Interaction*, 2001, 13(1): 41 ~ 51
- Cheng Y F. Skills and theory for Inputting Chinese Characters into computer with keyboard (in Chinese). *Chinese Institute of Information Beijing*, Beijing: China, 1996. 156 ~ 212  
(陈一凡. 汉字键盘输入技术与理论基础. 北京:北京信息学
- 院,1996. 156 ~ 212)
- Clive F, Richard H, Pam M. Recognition Accuracy and User Acceptance of Pen Interfaces Papers: Pen Interfaces. *Proceedings of ACM CHI '95 Conference on Human Factors in Computing Systems*. New York, USA: ACM Press/ Addison - Wesley Publishing Co, 1995, 1: 503 ~ 510
- Roy S K. VRUSE - a computerized diagnostic tool: for usability evaluation of virtual; synthetic environment systems. *Applied Ergonomics*, 1999, 30(1): 11 ~ 25
- Sherry P C, Robert D D, Beverly H W. The Effects of Recognition Accuracy and Vocabulary Size of a Speech Recognition System on Task Performance and User Acceptance Communication: Auditory and Vocal Communication. *Proceedings of the Human Factors Society 32nd Annual Meeting*, Santa Monica, CA: Human Factors Society, 1988, 1: 232 ~ 236
- I. Scott M. Fitts' Law as a Research and Design tool in Human Computer Interaction. *Human Computer Interaction*, 1992, 7: 91 ~ 139
- Doglas S. A. The Ergonomics of Computer Pointing Devices. London, UK: Springer - Verlag, 1997. 39 ~ 43
- Paul M. F. The Information Capacity of the Human Motor System in Controlling the Amplitude of Movement. *Journal of Experimental Psychology*, 1954, 47(6): 381 ~ 391
- Whitfield D., Ball R. G., Bird J. M. Some Comparison of On - display and Off - display Touch Input Devices for Interaction with Computer Generated Displays. *Ergonomics*, 1983, 26(11): 1033 ~ 1053
- Jellinek H. D, Card S. K, Powermice and User Performance Fitts Session [C]. *Proceedings of ACM CHI '90 Conference on Human Factors in Computing Systems*. New York, USA: ACM Press, 1990. 213 ~ 220
- Kantowitz B. H., Elvers G. G. Fitts' Law with an Isometric Controller: Effects of Order of Control and Control - Display Gain. *Journal of Motor Behavior*, 1988, 20(1): 53 ~ 56
- Graham E, L. MacKenzie C. Pointing on a Computer Display. *Proceedings of ACM CHI '95 Conference on Human Factors in Computing Systems*. New York, USA: ACM Press/ Addison - Wesley Publishing Co, 1995, Short Papers
- Zhang K, Chen Y F. A Cognitive Model of Typing Chinese into Computers. *Journal of Chinese Informatics*, 1992, 5(4): 13 ~ 19  
(张侃,陈一凡. 汉字键盘输入的认知模型. 中文信息学报, 1992, 5(4): 13 ~ 19)
- Gao D G, Zhong Y P. The Influence of the frequency of use on the speed of recognition in the Chinese characters (in Chinese). *Psychological Science*, 1995, 18: 225 ~ 228  
(高定国, 钟毅平. 字频影响汉字的认知速度的实验研究. 心理科学, 1995, 18: 225 ~ 228)
- John B E, Kieras D E. Using GOMS for User Interface Design and Evaluation: Which Techniques?. *ACM Transactions on Computer - Human Interaction*, 1996 3(4): 287 ~ 319
- Yang Z L. *Experimental Psychology (in Chinese)*. Shang Hai: China, East Normal University Press. 1996. 177  
(杨治良. 实验心理学. 上海:华东师范大学出版社,1996. 177)
- Wu Y S. Sun A X. *Probability Theory and Statistics (in Chinese)*. Guang Zhou: China, South China Tech. University Press, 1995. 180 ~

181 (吴亚森,孙爱霞. 概率论与数理统计. 广州:华南理工大学出版社,1995. 180~181)

20 Murata A. Empirical Evaluation of Performance Models of Pointing Accuracy and Speed with a PC Mouse. *International Journal of Human Computer Interaction*, 1996, 8(4): 457~469

21 Epps B. W. Comparison of Six Cursor Control Devices Based on Fitts' Law Models. *Proceedings of the 30th Annual Meeting of the Human Factors Society*, Santa Monica, CA: Human Factors Society, 1986. 327~331

22 John B E, Newell A. Cumulating the science of HCI: From S - R Compatibility to transcription typing. *Proceedings of ACM CHI '89 Conference on Human Factors in Computing Systems*. New York, USA: ACM Press, 1989. 109~114

23 Liu Y. Queuing Network Modeling of Human Performance of Concurrent Spatial and Verbal Tasks. *IEEE Transactions on Systems, Man, Cybernetics*, 1997. 27

24 Feyen R. Modeling Human Performance using the Queuing Network - Model Human Processor (QN - MHP). Doctor Thesis, University of Michigan, 2002

附 录

附录 1:在 N 个简单刺激中找到一个目标刺激的时间估计

实验心理学对反应时的研究发现,随着刺激数(被选项)的增加,个体的反应时出现一定规律的递减<sup>[18]</sup>:

刺激数/被选项 N	1	2	3	4	5	6	7	8	9	10
反应时(RT) (ms)	187	316	364	434	487	532	570	603	619	622

使用 SPSS 曲线拟合功能,对上述的递减数据定量描述为:  $RT = 172.766 + 199.037\ln(N)$ ; ( $R^2 = 0.991$ ,  $F(df = 8) = 865.91$ ,  $p < 0.001$ )

附录 2:手写笔选中随机出现在编辑区的错误识别汉字的平均移动距离的估计

设长方形编辑区(a × b)中每个汉字的尺寸为 1 个单位,编辑区长、宽的尺寸分别为 a、b 个单位(见图 4)。设笔从编辑区中心点 OE 出发,TO 表示编辑区中心点到各汉字的距离。

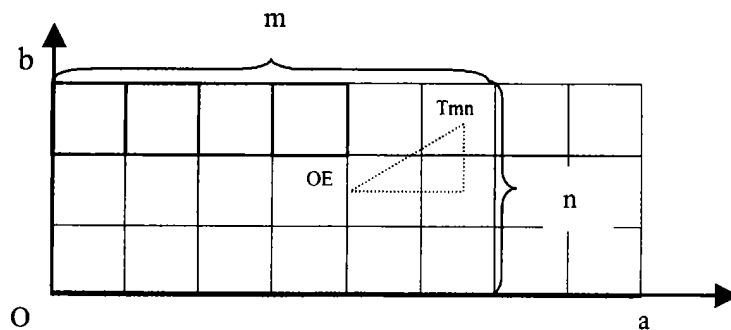


图 4 a × b 编辑区示意图

由上图,编辑区中心点 OE 的坐标为(0.5a, 0.5b); 第 n 行,第 m 列的汉字中心 Tmn 坐标为(m - 0.5, n - 0.5);

Tmn 与编辑区中心点 OE 的距离 TO 为:

$$[(0.5a - (m - 0.5))^2 + (0.5b - (n - 0.5))^2]^{0.5} = [(0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2]^{0.5}$$

Tmn 的分布密度函数为:

$$(x,y) = \begin{cases} 1/(ab) & x \in [0,a]; y \in [0,b] \\ 0, & \text{其它} \end{cases}$$

TO 的数学期望值:  $E(TO) = E([(0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2]^{0.5})$

$$= \int_{m=1}^a \int_{n=1}^b \{ [(0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2]^{0.5} \times [1/(ab)] \} = 1/(ab) \times \int_{m=1}^a \int_{n=1}^b [ (0.5a - m + 0.5)^2 + ($$



$$0.5b - n + 0.5)^2]^{0.5}$$

$$= 1/(ab) \times \left[ (0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2 \right]^{0.5} \left( \text{令 } \begin{matrix} a \\ m=1 \end{matrix} = \begin{matrix} a \\ m=1 \end{matrix}; \begin{matrix} b \\ n=1 \end{matrix} = \begin{matrix} b \\ n=1 \end{matrix} \right)$$

当  $a = 8, b = 3$  时,  $= > E(TO) = 2.209$

$D_{mef} = 12.8 + 94.7 \log_2(2A/W)$ ; (已设每个汉字的尺寸为 1 个单位  $= > W = 1$ )

另外,虽然 Fitts Law 是基于水平移动提出的,但 Astuo Murata (1996) 的研究发现,各种绩效模型在建立时可以不考虑运动的角度问题,关键是该模型本身与实际数据的拟合程度<sup>[20]</sup>。所以选中该汉字时间:  $D_{mef} = 12.8 + 94.7 \log_2\{2 \times 1/(ab) \times [(0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2]^{0.5}\}$  (13)

### 附录 3:各路线修改时间的一般式与在本研究中的具体估计值

路线	一般式	加入本研究已知界面参数后
路线 A	$TA = 612 + 199\ln(N) + 12.8 + 94.7 \log_2\{2 \times 1/(ab) \times \left[ (0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2 \right]^{0.5}\} + 612 + 199\ln(I) + 2 \times 12.8 + 2 \times 94.7 \log_2(2Lec/E(Wec)) + 12.8 + 94.7 \log_2\{2 \times 1/(ab) \times \left[ (0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2 \right]^{0.5}\}$	TA = 3002ms
路线 B	$TB = 612 + 199\ln(N) + 12.8 + 94.7 \log_2\{2 \times 1/(ab) \times \left[ (0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2 \right]^{0.5}\} + [612 + 199\ln(I)] + 2 \times [12.8 + 94.7 \log_2(2Leh/E(Weh))] + WT + R$	TB = 3339 + WT
路线 B1	$TB1 = 2 \times 439 + 2 \times [12.8 + 94.7 \log_2(2Leh/E(Weh))] + [612 + 199\ln(I)] + WT + R$	TB1 = 2757 + WT
路线 B2	$TB2 = 439 + [612 + 199\ln(I)] + 2 \times [12.8 + 94.7 \log_2(2Lec/E(Wec))] + 12.8 + 94.7 \log_2\{2 \times 1/(ab) \times \left[ (0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2 \right]^{0.5}\}$	TB2 = 1983ms
路线 B3	$TB3 = VD = 439ms$	TB3 = 439 ms

### 附录 4:修改时间 T 的一般式

$$T = n_1 TA + n_2 TB + n_3 TB1 + n_4 TB2 + [n_2 - n_3 - n_4] TB3 = n_1 (612 + 199\ln(N) + 12.8 + 94.7 \log_2\{2 \times 1/(ab) \times \left[ (0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2 \right]^{0.5}\} + 612 + 199\ln(I) + 2 \times 12.8 + 2 \times 94.7 \log_2(2Lec/E(Wec)) + 12.8 + 94.7 \log_2\{2 \times 1/(ab) \times \left[ (0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2 \right]^{0.5}\}) + n_2 (612 + 199\ln(N) + 12.8 + 94.7 \log_2\{2 \times 1/(ab) \times \left[ (0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2 \right]^{0.5}\} + [612 + 199\ln(I)] + 2 \times [12.8 + 94.7 \log_2(2Leh/E(Weh))] + WT + R) + n_3 (2 \times 439 + 2 \times [12.8 + 94.7 \log_2(2Leh/E(Weh))] + [612 + 199\ln(I)] + WT + R) + n_4 (439 + [612 + 199\ln(I)] + 2 \times [12.8 + 94.7 \log_2(2Lec/E(Wec))] + 12.8 + 94.7 \log_2\{2 \times 1/(ab) \times \left[ (0.5a - m + 0.5)^2 + (0.5b - n + 0.5)^2 \right]^{0.5}\}) + (n_2 - n_3 - n_4) \times 439 \quad (14)$$

手写绩效总时间的一般式:(2-13)式+(2-0)式

附录 5: 缩略字母表

缩略字母	中文释义	单位	缩略字母	中文释义	单位
ST	分割时间	ms	I	候选字个数(见图 1)	个
R	识别时间	ms	Lec	编辑区中心与候选区中心距离(见图 1)	mm
UD	用户延迟时间	ms	E(Wec)	编辑区与候选区平均宽度	mm
WT	手写速度	ms/字	Leh	编辑区中心与手写区中心距离(见图 1)	mm
RA	识别正确率	%	E(Weh)	编辑区与手写区平均宽度	mm
J	手写框平均边长(见图 2)	mm	Def	在编辑区定位错误识别汉字的时间	ms
MT	手写移动时间	ms	Dvcf	在候选区寻找正确候选字的时间	ms
N	正确输入的汉字的个数	个	Dec	从编辑区到候选区的移动时间	ms
a	编辑区长度(见图 4)	mm	Dmcf	在候选区定位正确候选字的时间	ms
b	编辑区宽度	mm	Dec	从候选区返回到编辑区的移动时间	ms
c	候选区长度	mm	Deh	从编辑区到手写区的移动时间	ms
d	候选区宽度	mm	D	任务完成时间	ms
F	手写输入框个数(见图 1)	个			

## HUMAN PERFORMANCE MODELING IN SPATIAL SEGMENTATION CHINESE HANDWRITING RECOGNIZER

Wu Changxu<sup>1</sup>, Li Huailong<sup>2</sup>, Yang Qunhui<sup>2</sup>, Zhang Kan<sup>2</sup>

<sup>(1)</sup> *University of Michigan, Ann Arbor, U. S. A. 48109*

<sup>(2)</sup> *Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China*

### Abstract

To predict and improve the human performance in spatial segmentation Chinese handwriting recognizer, a mathematic performance model was set up based on the Fitts' law and mathematical expectation of the discrete variables. The model was verified by the experiment that the model could account for the variance of human performance data in the experiment very well ( $R^2 > 0.85$ ) after the subjects experienced 8 times (8 × 24 characters) of training. The model can be applied into (1) the prediction of the task completion time when users are interacting with a specific spatial segmentation handwriting recognizer in copying task; (2) the choice of the optimal parameters in the recognizer; It was also found by the model that when the other parameters in the handwriting recognizer are constants, the reduction of user's handwriting time per character will play a more important role in determining human performance than the reduction of recognition time.

**Key words** mathematics psychological modeling, Chinese handwriting recognizer, human performance modeling, spatial segmentation.