

# 移动计算用户界面可用性评估的眼动方法

程时伟<sup>1</sup>, 石元伍<sup>2</sup>, 孙守迁<sup>1</sup>

(1. 浙江大学计算机科学与技术学院, 浙江杭州 310027; 2. 湖北工业大学艺术设计学院, 湖北武汉 430068)

**摘 要:** 为了降低传统可用性评估方法中的主观影响和误差, 本文应用眼动跟踪技术分析人机交互过程中的认知情况, 提出移动计算用户界面可用性评估方法. 该方法选取凝视时间、凝视数目、感兴趣区转移频次矩阵等眼动指标, 并结合交互任务与自顶向下的视觉认知机制对界面进行分析, 建立可用性评估模型, 考察视觉搜索和信息加工等认知效率, 进而为界面设计优化提供指导. 以手机用户界面为对象的评估实例表明, 该方法可有效分析界面认知内因, 获取可用性问题清单, 与用户主观评估保持较高一致性, 验证了其适用性与有效性.

**关键词:** 可用性评估; 眼动; 用户界面; 移动计算; 手机

**中图分类号:** TP391 **文献标识码:** A **文章编号:** 0372-2112 (2009) 4A-146-05

## An Approach to Usability Evaluation for Mobile Computing User Interface Based on Eye-tracking

CHENG Shi-wei<sup>1</sup>, SHI Yuan-wu<sup>2</sup>, SUN Shou-qian<sup>1</sup>

(1. School of Computer Science and Technology, Zhejiang University, Hangzhou, Zhejiang 310027, China;

2. School of Arts and Design, Hubei University of Technology, Wuhan, Hubei 430068, China)

**Abstract:** To reduce the subjective influences and errors from traditional usability evaluation, an approach based on eye-tracking was proposed for the mobile computing user interface. It analyzed the recognition issues in human-computer interaction, and selected fixation durations, numbers, matrix of fixation transfer frequency and other data related to eye movements as measuring criteria. Evaluation strategies were generated for efficiency of visual search and information processing, which explored the domain task and top-down visual mechanism to guide the optimization for interface design. The approach was applied for the usability evaluation with smart phones; as a result, it differentiated usability levels effectively and made consistency with subjective feedback. It was well done to illustrate the feasibility and effectiveness for the proposed approach.

**Key words:** usability evaluation; eye-tracking; user interface; mobile computing; smart phone

### 1 引言

由于手机、PDA 等移动计算设备外形和显示面积较小, 用户界面 (User Interface, UI) 应采取与 PC 软件系统不同的设计模式, 进行可用性评估可有效提升其设计水平, 增强产品市场竞争力. 传统的可用性评估方法主要包括启发式评估、焦点小组、认知走查、出声思考和绩效测试等<sup>[1~3]</sup>. 但这些方法存在一定的局限性: 前三种方法受被试用户主观影响较大, 且评估人员对被试用户也有一定的倾向引导, 导致评估结果缺乏客观性; 移动环境下 (如手机用户户外行走时), 出声法进一步增加了用户的认知负荷; 由于外界干扰, 实地的绩效测试很难控制实验变量, 且绩效指标的选取、数据记录和分析等依赖评估人员的知识和经验, 人为因素干扰较大.

与上述方法相比, 眼动跟踪 (Eye-tracking)<sup>[4,5]</sup> (简称

眼动) 采用高频摄像设备自动跟踪与记录用户眼球运动信息, 对用户干扰小, 且携带方便 (如头戴式眼动仪戴于用户头部, 可在室内外自由活动) 可以减少评估中的主观影响和实验误差, 且在一定程度上揭示用户内部认知差异, 提高评估效率与质量. 本文提出一种基于自顶向下 (Top-down) 视觉认知机制<sup>[6,7]</sup> 的评估方法, 通过设定眼动指标建立评估模型, 对用户界面进行认知内因分析. 手机评估实例表明, 该方法能有效提高可用性评估结果的客观性, 真实反映用户认知活动, 可为用户界面设计优化提供指导.

### 2 可用性评估的眼动方法

文献[8]结合眼动技术与传统绩效测试指标进行可用性评估, 但没有建立界面设计与眼动行为之间的映射关系, 评估结果很难用于设计改进. 本文则从交互任务

出发,根据自顶向下的视觉认知机制进行关键界面设计要素的眼动数据分析.具体的评估方法框架如图 1.

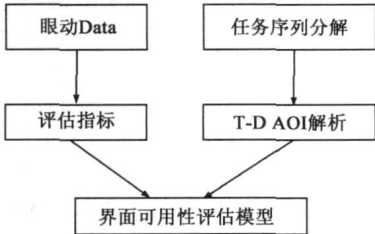


图1 眼动评估方法框架

评估模型主要对眼动数据进行参数化处理,提取有效的评估指标,并在 Top-down(T-D)视觉认知机制的基础上,结合交互任务序列建立认知活动与界面感兴趣区域(Area of Interest, AOI,即用户视觉注意的焦点区)之间的映射关系.

2.1 眼动数据

眼动行为在视觉认知过程中起着重要作用,它主要由凝视(fixation)和眼跳(saccade)交替出现构成序列<sup>[5]</sup>.凝视是指视线在界面上某一相对集中区域(视角0.5~1.3°)停留一定时间(通常至少在100~200ms以上),一般认为这种停顿主要用于获取信息进行认知处理.眼跳是指眼睛在两个凝视点之间的快速移动.

2.2 眼动评估指标

选取合理的指标可以简化计算复杂度,提高评估的信度、效度与效率.当前应用较多的指标包括凝视时间与凝视次数、眼跳数目与幅度、扫描路径、外接凸多边形、空间密度、击中目标率等<sup>[5]</sup>.本文选择手机UI作为研究对象,由于其屏幕尺寸较小,且受到眼动设备视角分辨率与采样精度,以及移动设备使用环境的限制,应避免选取对尺寸精度、定位效率等要求较高的指标.文献<sup>[9]</sup>给出了凝视时间、凝视数目、目标击中率等指标的定义与计算方法,本文在此基础上进行补充与完善,设定以下4个指标:

- 凝视持续时间( $t$ ).凝视可以用来表征用户进行信息获取或内部加工,用眼动仪记录每个凝视的持续时间.为了保证每个凝视均反映用户的有效认知情况,设定 $t$ 不小于100ms;
- 凝视数目( $n$ ).应用眼动仪在用户界面上记录的凝视点个数;
- 视觉捕获( $C$ ).将最初250ms<sup>[10]</sup>以内最先获得凝视的对象称为视觉捕获(Eye catch),用来描述最吸引用户注意的目标,一般通过查找眼动凝视数据记录的时间序列得到;
- AOI转移矩阵( $R$ ).计算完成一个交互任务的过程中,凝视在所选各AOI上转移频次得到的矩阵.例如一个3×3的AOI转移矩阵为:

$$R = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$a_{ij}$ 表示凝视从AOI <sub>$i$</sub> 转移到AOI <sub>$j$</sub> ( $i, j = 1, 2, 3$ )的频次.如 $a_{11}$ 表示前后两个凝视都在AOI<sub>1</sub>中的频次, $a_{12}$ 表示凝视从AOI<sub>1</sub>转移到AOI<sub>2</sub>的频次, $a_{21}$ 则表示凝视从AOI<sub>2</sub>转移到AOI<sub>1</sub>的频次.

2.3 任务序列分解

用户在使用移动计算设备时,绝大多数情况下都是有目的的执行一个特定任务,而每个任务都是一系

列的动作序列组成.只有用户正确执行了每个序列,才能保证成功完成任务.因此,为了跟踪用户在完成任务过程中的认知活动,就必须针对任务序列的每个环节进行分析.以

手机为例,给电话本里的某人打电话任务可分解成序列TSEQ= Seq1, Seq2, Seq3, Seq4, Seq5, Seq6,具体如表1定义:

2.4 自顶向下(Top-down)AOI解析

自顶向下的理论认为<sup>[6,7]</sup>视觉认知过程与任务相关,受意识控制,用户会根据任务目标和自身的经验、知识去主动搜索、注意感兴趣区域.根据该理论对眼动指标进行分析,就会避免那些底层视觉系统无意识地自底向上(Bottom-up)进行视觉选择而产生的噪声数据干扰.

结合上文对任务序列的分析,可以认为用户会主动进行选择 and 注视那些与任务执行相关的用户界面元素,如菜单、图标、按键等,可将这些设定为不同的AOI,并进一步与不同的认知活动(如视觉搜索与信息加工)联系起来.以表1的任务序列为例,可得表2所示的AOI解析:

| 表2 基于认知活动的AOI解析 |                         |      |
|-----------------|-------------------------|------|
| 任务序列            | 对应AOI                   | 认知活动 |
| Seq1            | AOI <sub>1</sub> :电话本菜单 | 搜索   |
| Seq2            | AOI <sub>2</sub> :电话本菜单 | 加工   |
| Seq3            | AOI <sub>3</sub> :搜索菜单  | 加工   |
| Seq4            | AOI <sub>4</sub> :联系人姓名 | 加工   |
| Seq5            | AOI <sub>5</sub> :联系人号码 | 搜索   |
| Seq6            | AOI <sub>6</sub> :联系人号码 | 加工   |

由于移动计算设备的屏幕较小,用户界面元素尺寸比普通PC上要小得多.另一方面,由于人眼有时对凝视周边的视觉对象也进行认知处理,就是用俗称的“余光”来感知对象.因此,如何在实际中合理定义解析

过的 AOI,并精确统计各 AOI 内的眼动数据,如凝视数目等,就显得非常重要.如图 2 所示.

判定一个凝视(虚线圆)是否为有效凝视,根据 AOI(实线矩形)中心与凝视中心的几何距离是否超过一个阈值  $k$  来决定,如图 2 中的凝视  $a$  位于 AOI 之内,是有效凝视;凝视  $b$  虽然超出 AOI 边界,但其与 AOI 的中心距  $d_1 < k$ ,因此也是有效凝视;而凝视  $c$  与 AOI 的中心距  $d_2 > k$ ,则其为无效凝视.于是有效感知区域即是以 AOI 中心为圆心,以  $k$  为半径的一个圆(图 2 中实线圆).如果手机屏幕大小为  $320 \times 240$  像素, $k$  通常取值为 10 像素,为了实际计算方便,通常将 AOI 的有效感知区域圆形的外切矩形(图 2 中虚线矩形)作为计算区域,若超出屏幕边界,或者与临近 AOI 重叠,将做相应裁剪.

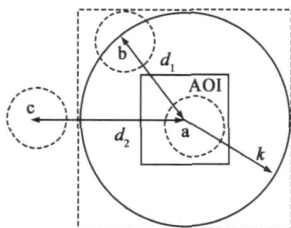


图2 有效凝视点计算

## 2.5 可用性评估模型

如何应用眼动指标有效反映用户的认知活动并发现用户界面设计的可用性问题,需要建立一个可行、有效的评估模型.

有关实验表明,同一个用户在不同的任务、场景、动机和目标下对同一认知对象的眼动指标也存在明显差异<sup>[11]</sup>.据此,本文主要针对手机用户在指定交互任务过程中的眼动行为进行研究,结合相应的认知活动与界面 AOI,提出评估模型 EM,其形式化为:

$$EM = TSEQ, \sum_{i=1}^m AOI_i(u, f), \bar{t}_i, n_i, AOI(C), R$$

其中, TSEQ 表示任务序列;  $\sum_{i=1}^m AOI_i(u, f)$  为包含  $m$  个 AOI 解析的集合,  $u$  为  $AOI_i$  所包含的界面元素(如某菜单项),  $f$  为该  $AOI_i$  所对应的认知活动类型(如搜索或加工);  $\bar{t}_i, n_i$ , 分别表示在  $AOI_i$  中的平均凝视持续时间和凝视数目;  $AOI(C)$  为包含视觉捕获对象的感兴趣区;  $R$  为该 AOI 集合的凝视转移频次矩阵.在该模型中:

•  $AOI_i$  中的凝视个数  $n_i$  与视觉搜索绩效相关,凝视过多表明用户界面设计不合理,导致视觉搜索效率降低;同时,  $n_i$  反映该 AOI 的重要和被关注程度,越重要的 AOI 一般  $n_i$  也越大;

•  $AOI_i$  中的总凝视持续时间  $\bar{t}_i \times n_i$  与加工效率有关,总凝视持续时间过多表明用户界面设计不合理,导致信息加工效率降低;同时也与该 AOI 的重要和被关注程度成正相关;

•  $AOI(C) = AOI_j (j = 1, 2, \dots, m)$ , 如果  $AOI_j$  不是任务序列中的首项或其他关键项,则说明对任务完成起决定作用的界面元素没有被用户发现或注意,则会降

低任务完成的效率或增加出错几率;

• 凝视转移频次矩阵  $R$  可反映用户的视觉扫描路径(scan path),若界面设计不合理,用户就会在相关 AOI 多次转移,反复搜索和加工,路径交叠错乱,非零值或取值大的矩阵项就会增多.

## 3 评估实例

评估实例以手机为对象,在文献[8,9]实验基础上,考察用户针对手机界面进行视觉搜索与信息加工的效率,从而衡量界面设计的可用性水平,为改进设计缺陷提供参考与指导.

### 3.1 实验设计

本文以索爱 W958c 音乐手机及其仿真原型作为实验材料,从而获得更全面与客观的评估效果.原型系统在交互流程、界面设计上都与真机保持高度一致,在 PC 上显示的分辨率  $240 \times 320$ ,用户利用鼠标模拟手机按键操作.原型测试虽与真机在操作模式、用户体验等方面有一定差异,但针对认知内因的效率评估仍然是有效的<sup>[11]</sup>,且能弥补真机测试受环境干扰较大的缺陷.

被测用户共 10 名,均为在校大学生,年龄介于 21 ~ 27 岁,男女各半.实验测试设备采用 ASL 公司的 Eye tracker R6 HS(桌面式,用于原型测试)和 Mobile eye(头戴式,用于真机测试)眼动仪,采样频率分别为 120Hz 与 60Hz,采样精度均为 0.5 视角(40cm 距离),如图 3 所示.

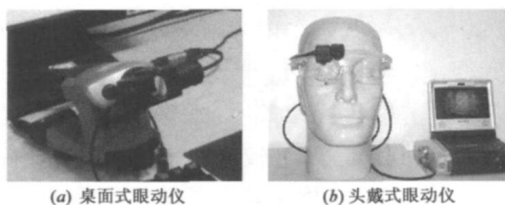


图3 眼动仪

### 3.2 实验过程

测试前用户阅读实验指导,熟悉任务后进行短试试用,然后开始正式测试.利用桌面式眼动仪进行原型测试,头戴式眼动仪进行真机测试,如图 4 所示.



图4 用户进行手机眼动可用性测试

实验完成后请用户填写调查问卷,获取主观评估与反馈.为消除学习效应,随机安排一半用户先进行原型测试,另一半先进行真机测试.

### 3.3 评估结果与讨论

限于篇幅,以其中一位用户的实验数据为例,该用

户的任务是找到音乐播放器的均衡器菜单,任务序列可表示为:

TESQ = {Seq1:搜索“设置”菜单},{Seq2 点击“设置”菜单},{Seq3:搜索“均衡器”菜单}.

相应的 AOI 解析为:

AOI<sub>1</sub>:(主菜单列表,搜索);

AOI<sub>2</sub>:(“设置”菜单项,加工);

AOI<sub>3</sub>:(弹出菜单列表,搜索);

AOI<sub>4</sub>:(“均衡器”菜单项,加工).

眼动数据如图 5 所示,虚线框即为各 AOI,黑色实心圆为凝视,上面标有该凝视的持续时间和先后序号.

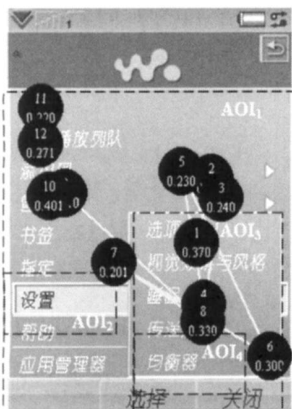


图5 原型测试眼动分析

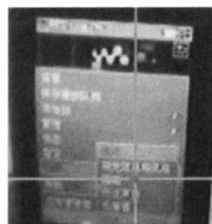


图6 真机测试眼动视频记录

各 AOI 平均凝视持续时间  $\bar{t}_1 = 0.276s$ ,  $\bar{t}_2 = 0.201s$ ,  $\bar{t}_3 = 0.280s$ ,  $\bar{t}_4 = 0.330s$ ;凝视数目分别为  $n_1 = 4$ (凝视 9 ~ 12),  $n_2 = 1$ (凝视 7),  $n_3 = 6$ (凝视 1 ~ 6),  $n_4 = 1$ (凝视 8);AOI(C) = AOI<sub>3</sub>,说明该用户能很快找到“设置”菜单并打开弹出菜单.此外,根据上文的定义可求得凝视转移频次矩阵为:

$$R = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 5 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$a_{33} = 5$ ,说明用户在弹出菜单列表对“均衡器”菜单项进行了较多搜索; $a_{41} = 1$ 说明用户找到“均衡器”后并不肯定,而是返回主菜单再次搜索,即  $a_{11} = 3$ .

真机测试的眼动视频记录如图 6 所示,十字准心即为当前凝视,表明用户也没有立即搜索到“均衡器”菜单项,搜索效率也较低.

据此得到可用性问题清单:

- 弹出列表菜单项过多,不利于用户搜索;
- “均衡器”语义不明确,放置于底部不易发现;
- “均衡器”菜单可放置于主菜单列表.

实验结束后采用 5 点等距量表让用户针对该任务进行主观满意度打分(1 ~ 5 分,5 最高).

从表 3 评分看出,用户对该界面设计的主观满意度

也不高,与眼动方法的评估结果具有较高一致性,验证了本文方法的有效性.

表 3 主观满意度评估结果

| 用户 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | M   | SD |
|----|---|---|---|---|---|---|---|---|---|----|-----|----|
| 得分 | 2 | 2 | 4 | 3 | 3 | 1 | 2 | 2 | 1 | 1  | 2.1 | 1  |

4 结束语与下一步工作

本文结合视觉认知和界面设计建立了基于眼动的可用性评估方法,针对移动计算设备可用性展开研究,评估实例表明该方法不仅能较有效评估用户的认知效率,还能得到可用性清单,如菜单列表过深等,可直接指导界面设计的改进.

由于眼动数据与内在认知心理活动之间的关系非常复杂,存在一定的系统误差,为提高其有效性与鲁棒性,下一步工作将结合肌电、脑电等测试手段,使评估方法更加全面与完善.

参考文献:

[1] 戴均开,葛列众.手机原型界面的可用性评估[J].人类工效学,2007,13(2):13-15+37.  
Dai Junkai, Ge Liezhong. Usability testing of mobile telephone's prototyping[J]. Chinese Journal of Ergonomics, 2007, 13(2): 13-15. (in Chinese)  
[2] Po S, Howard S, Vetere F, B Skov M. Heuristic evaluation and mobile usability: bridging the realism gap [A]. Proceedings of MobileHCI 2004, LNCS 3160 [C]. Berlin: Springer-Verlag, 2004. 49-60.  
[3] Guan Z, Lee S, Cuddihy E, Ramey J. The validity of the stimulated retrospective think-aloud method as measured by eye tracking [A]. Proceedings of the SIGCHI 2006 [C]. New York: ACM Press, 2006. 1253-1262.  
[4] 胡凤培,韩建立,葛列众.眼部跟踪和可用性测试研究综述[J].人类工效学,2005,11(2):52-54.  
Hu Fengpei, Han Jianli, Ge Liezhong. A survey of eye-tracking and usability testing[J]. Chinese Journal of Ergonomics, 2005, 11(2): 52-54. (in Chinese)  
[5] 张光强,沈模卫,陶嵘.可用性测试中的视线跟踪技术[J].人类工效学,2001,7(4):9-13.  
Zhang Guangqiang, Shen Mowei, Tao Rong. The application of eye tracking in usability tests [J]. Chinese Journal of Ergonomics, 2001, 7(4): 9-13. (in Chinese)  
[6] Lukander K. Mobile usability-measuring gaze point on handheld devices [D]. Helsinki: Helsinki University of Technology, 2003.  
[7] 张菁,沈兰荪, Feng D D. 基于视觉感知的图像检索的研究[J].电子学报,2008,36(3):494-499.  
Zhang Jing, Shen Lansun, Feng D D. A survey of image re-

trieval based on visual perception[J]. Acta Electronica Sinica, 2008, 36(3): 494 - 499. (in Chinese)

- [8] 安顺钰. 基于眼动追踪的手机界面可用性评估研究[D]. 杭州:浙江大学, 2008.

An Shunyu. Usability evaluation research on mobile phone user interface based on eye-tracking[D]. Hangzhou: Zhejiang University, 2008. (in Chinese)

- [9] 程时伟. 手机用户界面可用性评估的眼动模型[A]. 第四届和谐人机环境联合学术会议[C]. 北京:清华大学出版社, 2008. 380 - 385.

Cheng Shiwei. Usability evaluation model based on eye track for mobile UI[A]. Proceedings of the 4th Joint Conference on Harmonious Human Machine Environment [C]. Beijing: Tsinghua University Press, 2008. 380 - 385. (in Chinese)

- [10] Tzanidou E, Petre M, Minocha S, Grayson A. Combining eye tracking and conventional techniques indications of user-adaptability[A]. Proceedings of INTERACT 2005, LNCS 3585 [C]. Berlin: Springer-Verlag, 2005. 753 - 766.

- [11] Koivunen K, Kukkonen S, Lahtinen S, Rantala H, Sharmin S. Towards deeper understanding of how people perceive design in products[A]. In CADE2004 Web Proceedings of Computers in Art and Design Education Conference [C]. Sweden: Copenhagen Business School, Denmark & Malmö University, 2004.

#### 作者简介:



程时伟 男, 1981 年生于湖北省黄石市, 现为浙江大学计算机科学与技术学院在读博士生. 主要研究方向为人机交互、界面设计、可用性工程等.

E-mail: flishmen@sohu.com

石元伍 男, 1972 年生于湖北省武汉市, 现为湖北工业大学艺术设计学院副教授. 研究领域为数字化艺术与设计.

孙守迁 男, 1963 年生于浙江省杭州市, 现为浙江大学计算机科学与技术学院教授, 博士生导师. 主要研究领域为计算机辅助设计与概念设计、人机交互等.