

CONG YAN DONG DAO XING DONG
SHI JUE GEN ZONG YAN JIU XIN FANG FA

从眼动 到行动

视觉跟踪研究新方法

郑欢 编译

学林出版社

从眼动到行动

——视觉跟踪研究新方法

郑 欢 编译

学林出版社

图书在版编目(CIP)数据

从眼动到行动:视觉跟踪新方法/郑欢编译.

—上海:学林出版社,2010.7

ISBN 978-7-5486-0035-0

I. ①从… II. ①郑… III. ①计算机视觉-视觉跟踪
IV. ①TP391.41

中国版本图书馆 CIP 数据核字(2010)第 108390 号

从眼动到行动——视觉跟踪研究新方法



编 译——郑 欢

特约编辑——杨晓青

责任编辑——吴耀根

封面设计——包雷晶

出 版——上海世纪出版股份有限公司

学林出版社(上海钦州南路 81 号 3 楼)

电话:64515005 传真:64515005

发 行——新华书店上海发行所

学林图书发行部(钦州南路 81 号 1 楼)

电话:64515012 传真:64844088

照 排——南京展望文化发展有限公司

印 刷——上海展强印刷有限公司

开 本——890×1240 1/32

印 张——7.25

字 数——20 万

版 次——2010 年 7 月第 1 版

2010 年 7 月第 1 次印刷

书 号——ISBN 978-7-5486-0035-0/J·3

定 价——20.00 元

(如发生印刷、装订质量问题,读者可向工厂调换。)

目录

行为研究——不仅是对专业工具的挑战	1
基于模型的眼动研究方法	13
基于语言和面部识别的宫泽贤治虚拟认知模式	31
看上去很高兴：积极的视觉注意力偏移的实验性控制	46
阅读中的眼球运动：新发现	58
使用眼动技术重现感觉图像	66
专业领域知识对图像再认过程中视觉热点的弱化影响	73
前期知识对查看和转译宏观表征和微观表征图像的影响	100
为繁杂视觉效果选择色标的感识处理方法	121
关于广告注意力的目标控制——亚尔布斯(Yarbus)的推论	139
特写广告关注的销售效果：贝叶斯中介分析	159
计量生物学和眼动轨迹视角——快进浏览中的电视广告信息 处理	184
态度和社会认知：社会排他性和早期人际知觉 ——选择性注意认同的信号	202
后记	222

行为研究——不仅是对专业工具的挑战

Proceeding studies on behavior — not only a challenge for professional tools

帕斯卡尔·曼戈尔德 *

Pascal Mangold

摘要：本文的观点均基于公司长期以来在行为研究领域系统开发实验的基础上。本文就日常行为研究的数据采集和分析进行了多方面的探讨，它指出运用专业软件进行行为研究的必要性，并表明音像资料不仅可为文档储存有益，同时还有利于分析。本文还探讨了使用结构化编码较之摘录笔记的优势。最后是针对互动软件工具环境的概述。

一、使用软件工具进行行为学研究是明智的

若依照以下步骤，我们会发现行为学研究看似蛮简单。

* 帕斯卡尔·曼戈尔德拥有信息科学文凭，并自1989年以来一直活跃在IT行业。他有自己的公司，开发科学和工业应用方面的IT系统，侧重于行为研究。他亦是一位作家，负责大学的培训研讨会、讲座，同时，他还是位广受欢迎的工业顾问。Graf von Deym Str. 5, D - 94424 Arnstorf, pascal.mangold@behavioral-research.com.

文章来源：《电路与系统学报》，第十期，第五卷，第1536~1543页，2006年10月。

WSEAS Transactions on Circuits and Systems, issue 10, vol. 5, pp. 1536 - 1543, October 2006.

我们“唯一”要做的事就是依据(图1)完成上述步骤。但是实际情况并非如此:过去十年里,我们在这个领域碰到的人不是电气工程师、信息学家就是计算机专家,他们也可能是心理学家,设计师、社会学家或其他类似的非专业技术工作人员。对他们而言做这样的行为研究远非易事。许多问题需要在测试前、测试中和测试后完成:

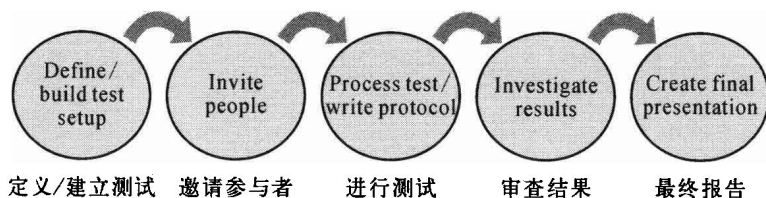


图1 行为研究的主要过程

如何巧妙地设定场景?如何对收集的数据进行进一步处理?数据将如何被集成?应该采取什么样的分析过程,且涉及到谁?等等。

如果影片存储在古典录影带里,那么重看这些影碟可能会受到限制。几个人可能无法在办公室看相同或不同的影片。因此,视频分析遇到一个严重瓶颈。哪些工具(软件/纸和铅笔)应该用于收集结果/译码/改编和以何种形式(带/不带时间戳,是基于绝对的时间或只是相对的时间,是结构化的列出还是以自由本文的形式注释等)?基于停滞时间的精确度而得出的数据是否会令人满意(而谁又该负责处理暂停时间)或在视频播放器装机械计时器?

如何将这些数据放在一个团队中共享(是每个人都享用同一文件还是拥有自己的副本)?如何才控制并集成结果(图表,统计的数字,任何种专业的可视系统)?

以上所列零星的项目表明,行为研究不只是(图1)列出的那么简单。它说明,最多出现的问题都不是在主要阶段产生的(如图1:做了1小时录像或书面协议),但在连续各个阶段完成任务之间还有很多事情必须完成(如将各种数据源整合,格式化时间尺度,以结构化方式收

集结果等)。图 2 更好地、更逼真地向我们展示了这一现实情况。

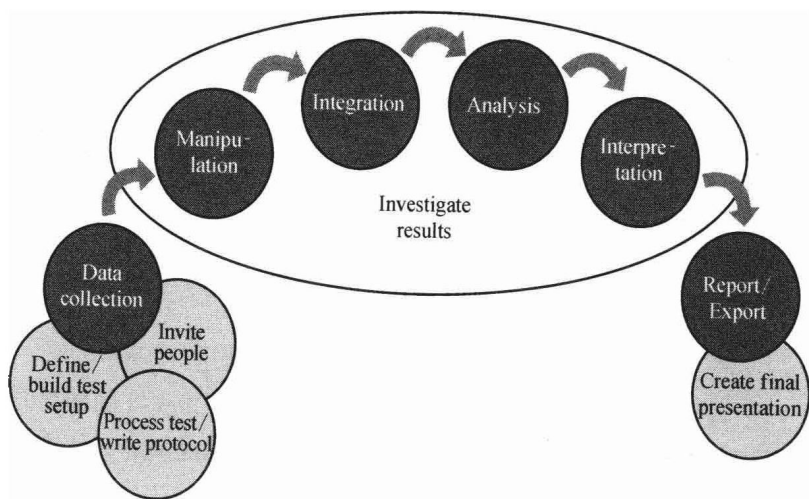


图 2 “真正的”行为研究过程

图 2 还给出了如果测试设施错误,信息丢失或被意外重组等状况的预测。效果如图 3 所示。最坏的情况是,人们发现结果不可靠

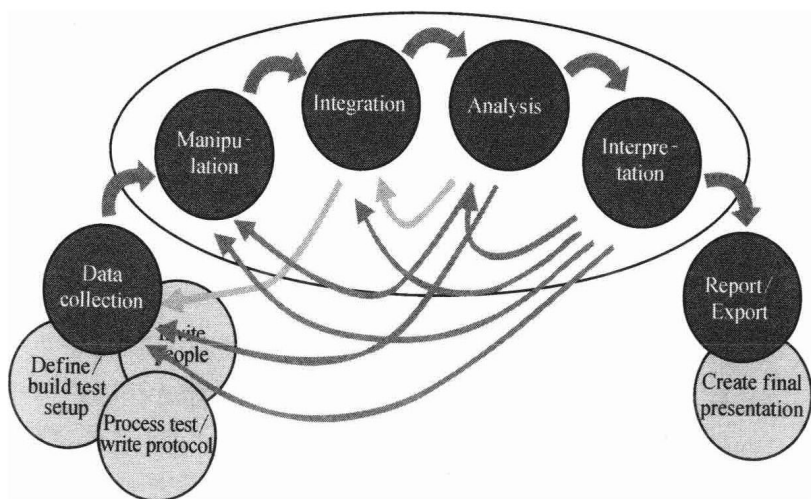


图 3 行为研究中糟糕的状况

或好奇：那样你不得不重新开始，不仅是“收集一些数据”而是把图 1 所示的每个步骤再进行一次(当然这问题到底有多大)。这需要耗费你很多时间而且确实是让人头痛的事情。

由于上述原因,我们的主张,“行为研究”领域(无论它的载体是可用性测试,心理学研究,人机工程学研究等),需要有专门的工具。通过使用这些工具,在本章开头部分讨论过的很多问题都可以避免,至少可以把它减少到最低限度。一个所谓的“工具链”(工具,“理解”彼此没有太多的用户直接干预)可以帮助做到:

- 用简单、无差错和结构化的方法收集数据
- 在任何时间操控数据,以便它与另一个数据进行匹配
- 轻轻用鼠标一点就可以将结果和原始数据相关联(例如显示视频中我的立场和生理数据使之与人类活动观察家提供的具体意见相匹配)

- 与其他程序交换和输出数据

这有助于节省时间且,最重要的是,提高搜索结果的质量。

1. 结构化的数据收集和结果的质量

即使不使用工具,人们也可以创造有趣的结果。只需通过简单地观察现场并用纸和笔写下评论/结果即可。这是一种“专家们”常见的做法。当然在此前他们已经通过长期的经验获得了关于“正确”或“错误”的潜意识性知识(一般意义上)。专家们常常无法正式地解释他们的发现。这样的结果就产生了一个重要的问题:专家确信的结果如何被鉴定其质量?

我们必须用一种方法来“科学地”实现至少以下标准:

- 客观性: 研究结果必须经受再省
- 重复性: 在相同情况下重复试验得出的结果必须相同
- 易于理解性: 寻找结论的方式以及结论本身都易于理解

举例说明: 一项关于网络的可用性的研究,9 个“可用性专家”得出的结论就完全不一。这些专家们一共发现了 174 项完全不同的可用性问题。61%的问题只有他们中的某一个人提到。只有 18%的问

题被至少三分之一的专家们提出。这也就意味着这样一个事实：平均每位专家只指出了所有相关结果中的 31% 而已（最高为 61%）。这是一个应该引起警示的问题。

同时我们还看到单凭“专家知识”不足以使客观性和易于理解性做到令人满意。因此，仅仅依靠“专家知识”来测试结果远非易事。

其次，凭借专家还导致了另一个问题，就是如何把这些知识转录，共享并与其他人的知识合并以提高和创造新的见解？

这些问题使结果正式化和结构化就显得尤为重要。

2. 做笔记与转录编码

在所有方法中，以下数据收集方法被广泛用于观察：

- 记笔记：观察员将他/她认为有趣的东西记录下来。这些记录可以是在未经解释的“原始数据”（如仅仅是一个人说的话）和已经找到结论之间的任何事情。

- 转录：此方法是用于将人们所说和所发生的事情，用一种明确的方式记录下来（如[呼气]我不明白这个[感叹][~]噢，我的天哪！）。可以想象，以此来看，转录会变得非常复杂。

- 编码：以标示符或者数字（已经附加到某些时间点或者时段的数字）来定义的标签。“编码模式”（也常被称为“分类模式”）主要是分层，当然也可以变得非常复杂。如某人谈论颜色的积极作用：[表达] - [口头] / [2001] - [颜色] / [品质] - [积极]……

所有这些方法都各有利弊。其中一些对于行为的分析，列于下表：

方 法	优 点	缺 点
做笔记	速度较快	由于缺乏结构性，笔记往往记得并不清晰，使得后期分析“不得其意” 笔记的最初稿并不易于理解也不客观

续 表

方 法	优 点	缺 点
转录	如果这一步骤做得有足够的条理结构性,那么进一步分析还有可能 适用于文本的上下文分析 客观 易于理解	耗时 如果最后反映的只是我们可以从视频中得知的,那么就是浪费时间 不算是行为研究中最好的方法 需要大量的研究来确定一个很好的抄写文本
编码	速度非常快 后续自动研究成为可能 客观 易于理解	需要大量的研究来确定一个好的编码方案

结构之间的真正转录和编码的差异并不太大。但在常规理念里,用一种或多或少有条理结构的方式记录事情要胜于开发一个好的编码方案。从我们的经验看来,建立一个编码方案总是可能的!至少先有一个粗略的,然后在研究过程中去完善它。我们已经看到这样的学生,像奴隶一样地疲于记下他们从视频磁带中所看到或者听到的。只为让主管得到 MS Word 格式的文件,而这些反映到文字不过是实际记录中的一个小片段。其次,直到这个进程的现阶段,对发生的一切没有任何解释。

以上经验显示了编码的优点:

- 快捷: 通常分析 X 小时的摄录需要 X 小时。因此时间是个至关重要的因素!

- 易于理解: 要培养成功的程序员,必须让大家明白什么时候给什么样的代码。

- 客观: 为了保证质量,证明程序员有相同的共识且在同一时间看到同样事情的统计是必不可少的。

- 自动化可能性: 分析大型数据集以及部分数据的比较需要自动化。这样可以防止手工运算中的错误,同时可以节省很多时间。

二、现场观察与视频分析

试图在没有工具支持的情况下编码一个录像带是非常困难和耗费时间的东西。程序快捷键和质量实时监控丢失,使得编码过程中存在某种不自信。至少,结果很难被证明。这就是为什么很多人试着脱离视频而工作。

人类的认知是有限的,因此,观察任何事情意味着一次只能发现“什么是真正发生”的一个小片段,不同的人在同一时间观察同一场景有不同的重点。这意味着在观察中:

- 大量的信息未被觉察
- 你不知道你是否错过了一个有趣的或没趣的东西
- 你不知道你是否把注意力放对了地方

为了防止不堪重负的信息,很多信息被认知系统人为的过滤掉了。但是,不只是这部分信息缺失是问题所在,由于信息不全导致的错误理解才是更为严重的问题。如果在观察者的视线内,很多事情快速发生,而同时观察员在巨大压力下同步写下调查结果,笔记摘要很有可能根本不能正确反映事情的发生。这种情况还会因观察员在观察阶段试图修改之前笔记的常规行为而变得更糟。这样,他们就丢失了更多的信息,于是也就丢失了他们跟踪的事态。从这几年来这个领域的不同问题上给过的咨询来看,以上所述的行为都被观察得知是常规行为。我不是说现场观察完全无用,通过使用简单的观察计划,现场观察可能是有助和可靠的。但是在更复杂的情况下,要得到可靠的结果就很困难了。现场观察,显然不能完全符合“科学方法”的标准。至少没有一个特定的测试可以准确重复。

如果一个方法不具备重复性,那么也就是说随后可以尝试其他获得可信度的方法:大量的素材要被处理以获得统计学意义,从而得到被认为可以信赖的结果。现在,这里有个实际问题:除总成本总是一个限制因素之外,更为困难的是找到大量的测试对象以及完

成这些测试的足够时间。为了解决这个问题,我们通过实验确定一个典型性行为:“我们摄录我们测试的所有会议”,“但是我们永远不会再看它们了,这太耗时。”我相信,几乎所有做过视频分析却没有使用专业视频分析工具方法的人都会同意以上观点。

但使用录像是行为研究的绝好办法:你可以重放、慢放和暂停“现实重现”或任何你想要的方式。让你可以作注解和编码,调查细节,随时修改收集到的资料,任意速度重看录像。你可以在几轮播放过程中重点看行为的不同方面,而免于超载的信息和任务。除了上述提到的重复,视频材料也可以用来实现可理解性,结果可视性。

在下面的内容中,我们将阐述如何使用工具支持视频分析,以尽量减少时间的花费并最大限度得出正确的结论。

三、视频分析工具

如今市场上已经有了视频分析软件。但依照我们的经验,真正有价值的行为分析软件需要综合以下几个不同方面:

- 定性 + 定量
- 预定义 + 探究
- 结构 + 免费注释

除了上述要点,开发和使用编码与注释,数据结构化和免费的结合同样也是必不可少的。

在评估了一些录像分析的商用软件系统后:“通过探测互动软件系统在复制未来设计研究方面的功能和特点,我们认为它完全满足了我们的要求。”因此互动的使用被证明是有效的,省时并欲提供一个显著地优势,用来分析和加强研究进程的效率。

有了互动软件,用户可以在分析过程的任何时间,采取不同的方式收集数据:

- 简单地用一个键输入预定代码
- 输入任一种文本的新代码

- 输入任何文字注释

此信息总是存储在时间开始和结束的节点之间。因此,随时都可以给出确切的相关数据和视频。由于其时间格式使用的是所谓的标准时间码(小时:分钟:秒:视频帧),这就可以识别视频中每个单张图片以及必不可少的情感研究,面部表情,手势和姿势。互动软件用户的图片如图4所示,显示了其数据收集过程。

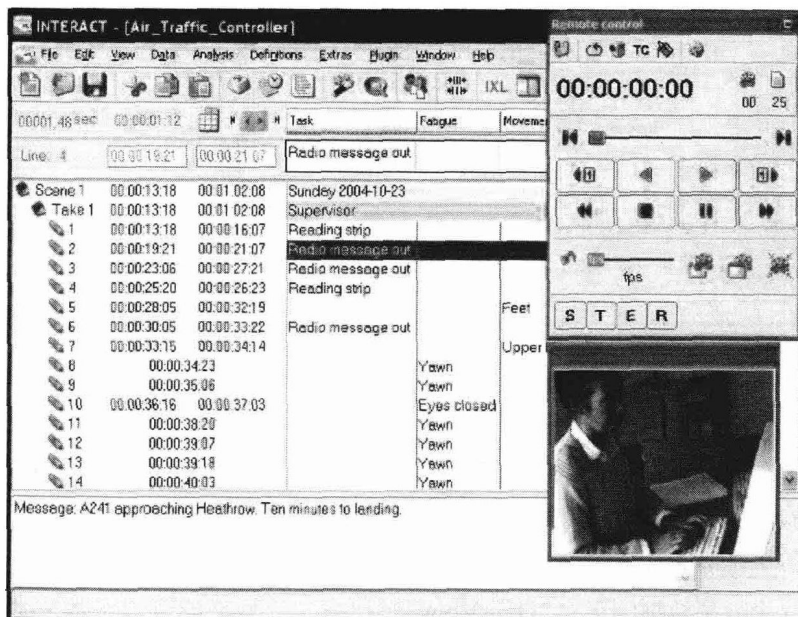


图4 用户互动界面

互动软件能够同步处理无限视频信号源并予以控制。这是必要的,如果有多个视频源被同时记录,那么没有必要把他们混合进一个视频,或者说不同视频的大小和质量应被分别单独保存。

互动软件提供了一个所谓的技术插件允许用户添加任何功能,虽然这一技术还未被集成。因此,互动软件工具几乎具有无限的可能性。所收集的数据不仅可以以任何目的被输出,同时其他类型的数据也可以输入进互动系统。

互动软件提供了大量的分析功能。其中,“信度”可使计算方便,提供了国际认可的数据质量测算;而序列分析表明,哪些代码紧随哪些代码,且有什么统计学意义。统计资料,例如频率,持续时间,百分比,方差和标准差也在整合中。图5为编码数据的特殊的时点刻线图。

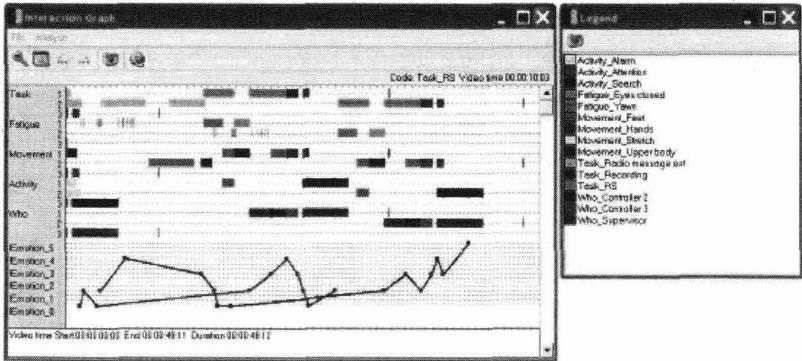


图5 时间编码数据线可视化

我们把互动软件中收集的数据称作“稀疏”数据(不定时的某些事件)。与之相对应的,如果在短时间段内(如生理数据,每十秒记录一次)集成的具有较高价值的数据,我们则称之为“高频率”数据。导入或整合的高频数据到互动软件不会有太多意思。那就是为什么一个分离数据的工具“DateView”被研发。这允许的可视化数字数据可存储在一个没有上线的图表里。此软件工具会自动与互动软件同步运行。在互动软件下,每当视频移动到另一个位置,数据图表正随之移动。

图6展示了桌面三个视屏窗口(被测试的人,他的环境和他的屏幕内容)的复杂分析印象。若干生物数据可视化和时间积编码。屏幕为 1400×1050 像素的分辨率,使用宏基 TravelMate 6000 笔记本电脑分析。正如人们所能看到的,那并没有足够的空间把所有东西都看得很清楚。特别是原始全屏下(1024×786 浏览)视频需要进

行审查,否则图像变得太模糊。因此,我们强烈建议在分析过程中另加一台显示器。

这些数据收集无限的可能性,同步音频无限可视的实际操作性,视频和数据文件使得互动软件和它的附加工具成为行为研究应用中的标准。

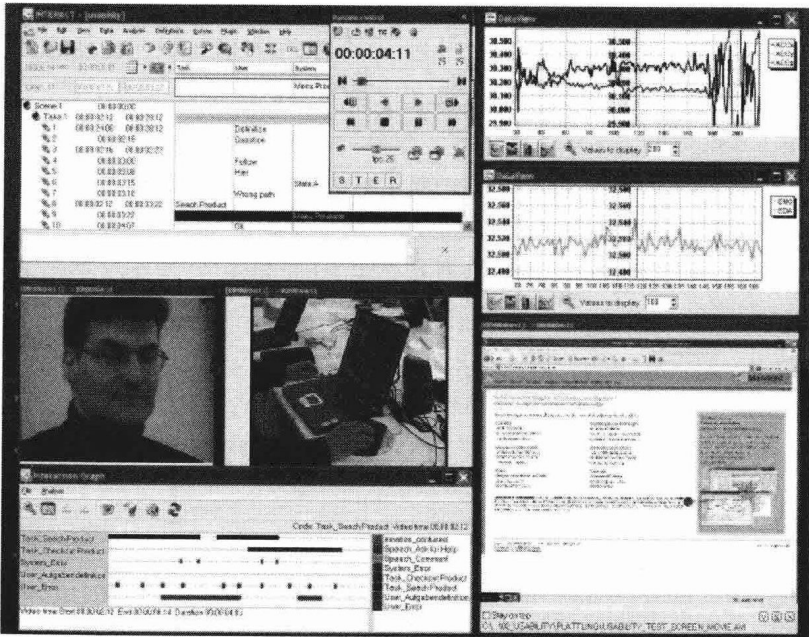


图6 被测试人桌面三个视屏窗口分析

那种对于某项任务有充分时间的感觉不可能存在于真实的日常工作中。如果人们觉得有充足的时间,那他们大概是:

- 1) 尝试把更多的任务放在当下而忽略最初的要求。
- 2) 简单地停止任务,当它们要“做他们的工作”的时候。

现在来看看,在情况1)中的人根本没有时间去顾及结果的质量,因为他们自我生产任务/数据/结果已经超负荷,而忙做一团。这和那些没有足够时间的人的结果是一样的。而情况2)的人因为没有什么剩余时间,或者即使有些许剩余时间,也在忙于“做他们的

工作”。

从所有上述所写和我们的经验来看,我们知道,使用工具肯定有助于节省大量的时间。但我们也知道,这里所指的时间是没有空闲时间。因为马上又有新的任务跟上。这也是为什么用户们即使用了专业工具,也从来没有过节省时间的感觉——除非在使用工具之前他们做过同样的事情。但这往往不会发生:我们发现那些使用工具的人说,“他更容易也更有趣了——但不怎么节省时间”——是的,确实如此,因为他们在做的完全是另一些事情。主要的任务不再占用他主要的时间了。那也是为什么有充足的时间做其他事情。

我们的结论是,时间因素绝对是对结果的质量有影响。

- 如果有较少的时间,质量几乎总是受到影响。尤其是在某些研究中,高度精准度的编码是很有必要的(面部表情,情绪状态检测,观察互动小组讨论的几个对象,确定后续的事情/有彼此相互依存)。

- 如果有(或将有)大量时间用来提高质量,则是理想情况。同时还需要这段时间所做的事情对原始任务没有任何影响。

我的结论就是,使用专业工具可以使投资得到很好的回报,尽管这个回报不是立竿见影的。

我们所定义的始终是“软影响”,如果用户之前做过视频手动分析,这种减压对项目中的其他因素有着积极影响。一些情况在没有工具的情况下不可能做到,这也是事实。

无论如何,衡量投资的回报是非常困难的,同时是强烈依据视频分析和工具支持情况而定的。

基于模型的眼动研究方法

A model-based approach to video-based eye tracking

李锋, 苏珊芒恩和杰夫皮尔兹*

Feng Li, Susan Munn and Jeff Pelz

摘要: 基于视频的眼睛跟踪系统通常是指跟踪瞳孔和一种表面依赖角膜反射(CR)的照明光源。在这两个功能之间的不同位置是用来确定观察者眼睛的方向。随着头部使用安装眼跟踪器,这种定位难免受限于眼睛与摄像机之间的相对运动,当事人视觉生理体征等的影响。基于视频跟踪器产生的问题也受到角膜反射检测的影响,如虚假的反射被误以为是合适的角膜反射。我们的方法是通过将这些问题模型化,然后探讨这些功能(瞳孔与角膜反射)是如何受到眼睛的相对运动的影响。光的关系之间的偏移量明显存在于瞳孔的中心,我们5个观察员进行了一项试验,以支持这些推导。在解决方案中使用一些解决办法用以解决上述的情况,在减少眼睛最后方向的数据中,噪声第一次用以弥补在摄像头方面眼睛所进行的跟踪

* 作者生活在美国罗切斯特,就职于罗切斯特理工技术学院,切斯特卡尔逊楼成像科学中心

文章来源:《现代光学杂志》第55期, Nos. 4-5, 2008年2月20日—3月10日, 第503—531页。

Journal of Modern Optics, Vol. 55, Nos. 4—5, 20 February - 10 March 2008, 503—531;

EBSCO 心理学数据库。