

虚拟现实人机交互的 eGOMS 模型研究

王 晖¹ 戴国忠¹ 袁宇明² 张凤军¹ 王丹力¹ 傅小兰²

(1 中国科学院 软件研究所, 北京 100080;

2 中国科学院 心理研究所, 北京 100101)

摘要: 采用心理学和计算机技术相交叉的研究方法, 对人机交互中人的认知过程的人类视觉特征和心理加工机制与特点等进行了细致的探讨; 结合虚拟环境中人机交互过程的不确定性、连续性和实时反馈性以及交互可用性评价指标的多样性等特点, 提出了适用于虚拟现实人机交互的扩展 GOMS(eGOMS) 模型, 克服了传统 GOMS 模型要求精确的用户操作、交互界面与实际应用相分离及评价指标单一的不足. 以 eGOMS 模型为基础对 PIBG 交互范式进行了评估, PIBG 为 VR 用户界面提供自然、隐式交互的手段, 从而大大降低了用户在虚拟环境下交互的认知负荷.

关键词: 虚拟现实; 人机交互; GOMS 模型



中图分类号: TP311; TP391

文献标识码: A

文章编号: 1671-4512(2005)S1-0322-04

Study of eGOMS model for human computer interaction in VR

Wang Hui Dai Guozhong Xuan Yuming Zhang Fengjun Wang Danli Fu Xiaolan

Abstract: This paper studies the mechanism and characteristics human perception and cognition, attention and the process of information, and proposes an eGOMS model for HCI in VR. User Model based on cognitive psychology is the base of interface design for Virtual Reality (VR). The research method of this paper is the combination of physiology and computer technologies. eGOMS overcomes the disadvantages of the traditional GOMS model, such as accurate operation, division of interface and applications, and mono criterion. PIBG interaction was evaluated with eGOMS, and the results show that PIBG allow users to interact with a computer freely through natural interface, which enhances the efficiency of interaction between users and computers, and reduces the cognitive burden of users in VR environment.

Key words: virtual reality (VR); human-computer interaction (HCI); GOMS model

Wang Hui Assoc. Prof.; Institute of Software, Chinese Academy of Science, Beijing 100080, China.

用户界面的发展到现在经历了三个主要的时代. 批处理界面、命令行界面和图形用户界面分别代表了三个时代中主流的用户界面. 到目前为止, 图形用户界面仍然是占统治地位的一类界面. 这种基于桌面隐喻、使用 WIMP 范式的界面之所以能够成为近二十年中占统治地位的界面, 是因为它与之前的界面相比, 具有对象可视化、语法极小化和快速语义反馈等非常明显的优点. 但随着计算机硬件设备的进步和软件技术的发展, WIMP 界面的缺点逐渐地体现出来. 目前研究者们将研究的焦点聚集到下一代的用户界面上, 提出了

Post-WIMP(或 Non-WIMP)的界面形式^[1].

虚拟现实系统向用户提供临境和多感觉通道体验, 它的三个重要特点即临境感、交互性、构想性^[2]决定了它与以往人机交互技术的不同特点. 在传统的人机系统中, 人是操作者, 机器只是被动地反应; 在一般的计算机系统中, 人是用户, 人与计算机之间以一种对话方式工作; 在虚拟现实中, 人是主动参与者, 复杂系统中可能有许多参与者共同在以计算机网络系统为基础的虚拟环境中协同工作. 作为一种新型人机交互形式, 虚拟现实技术比以前人机交互形式都有希望彻底实现和谐

收稿日期: 2005-08-13.

作者简介: 王 晖(1967-), 男, 副研究员; 北京, 中国科学院软件研究所(100080).

E-mail: wanghui@ios.cn

基金项目: 国家重点基础研究发展计划资助项目(9732002CB312103); 国家自然科学基金资助项目(60373056, 60303019).

的、“以人为中心”的人机界面.与传统的用户界面相比,虚拟现实中的人机交互有很多新的特点,如对象图形、对象种类、对象行为、通信方式等^[3].

本文对人机交互中人的认知过程进行了细致的探讨,包括人类视知觉特征和心理加工机制,从认知心理学的角度研究单通道和多通道信息加工的机制,包括人对单通道和多通道感知信息的加工、注意选择与分配的机制与特点,以及多通道信息整合过程进行了细致地研究;并在此基础上对 GOMS 模型进行了扩展,形成了 eGOMS 模型,并对 PIBG 交互范式进行了评估.

1 人机交互和认知加工过程

从信息传递的角度来看,人与计算机的交互主要可分为两种信息传递通路:一方面是计算机向用户提供信息,这是计算机的信息输出,人的信息输入;另一方面是用户向计算机提供信息,这是计算机的信息输入,人的信息输出.要建立更自然、高效的虚拟现实用户界面,需要充分利用不同认知层次和多个通道的信息,改善信息的输入和输出效率.

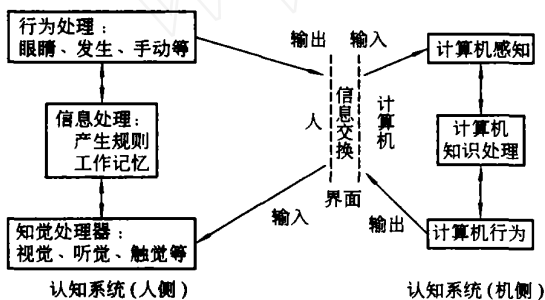


图1 人机交互界面和认知加工过程

从计算机输出来看,可分为二个层次,一是低级的感知觉加工过程,例如对颜色、声音的感知,多个通道的信息输入进来以后如何进行整合的感知觉加工,以及如何控制注意;二是高级的认知过程,主要涉及到词汇信息、语义和语法知识的加工.这两种心理过程从低到高逐步进行,并且高级认知过程可以反过来影响低级加工过程.

感知觉层次的加工机制.计算机提供的信息首先要经过感觉系统的加工,然后由知觉过程对感觉信息进行组织和解释,才能获得对感觉信息的意义.在这个过程中,计算机提供信息的方式会对用户的感知觉过程产生很大的影响,因此了解计算机的信息呈现方式与用户认知过程之间的关系对改善界面信息呈现方式,提高用户对计算机信息快速有效的感知觉有着很重要的影响.感知

觉层次的加工机制研究的一个重要目的就是减少用户的认知负荷.

认知负荷可以分为外在认知负荷和有效认知负荷.外在认知负荷能够干扰用户与界面的交互.有效的认知负荷与外在认知负荷正好相反,可以促进用户与界面的交互,导致图式获得和自动化.用户与界面交互时所需的所有认知负荷不能超越用户的认知资源.出现认知过载的一个因素是因为人类大脑的工作记忆容量的特点.文献[4]提出了一个工作记忆模型,模型中视觉信息和语义信息的存储是彼此独立,而且其存储容量是非常有限的.其中视觉工作记忆只能记忆3~4个单元(特征或物体),语义材料记忆也仅7个单元左右.

另一个影响认知负荷的关键因素是注意.根据认知心理学理论,外界的感觉信息是有限的,但其中只有有限的信息能进入中枢系统得到进一步加工,而没被注意的信息就会被过滤掉或者逐渐衰减.因此,注意在人的感知和认知加工中起重要作用.认知心理学研究表明,对人的认知加工来说,在无限的感觉信息输入和无限的长时记忆容量之间存在的瓶颈是工作记忆容量和注意资源的有限性.提供对认知心理过程的研究,为建立虚拟环境下人机交互模型奠定了基础.

2 虚拟现实人机交互的 eGOMS 模型

文献[5]提出了两个重要模型,即:人类处理机模型和 GOMS 模型.其中 GOMS 模型主要用于指导第一代(命令行)和第二代(WIMP)人机交互界面的设计和评价.GOMS 模型是关于用户在与系统交互时使用的知识和认知过程的模型,GOMS 是一个缩略语,代表目标(Goals)、操作(Operations)、方法(Methods)、选择规则(Selection rules):

“目标”指的是用户要达到什么目的,如查找某个网站.

“操作”指的是为了达到目标而使用的认知过程和物理行为(如先搜索引擎,再思考关键字,然后在搜索引擎中输入关键字).“目标”和“操作”不同,分别指“达到”某个目标和“执行”某个操作.

“方法”是指为了达到目标而采用的具体步骤(如使用鼠标单击输入域,输入关键字,再单击“查找”按钮).

“选择规则”用于选择具体方法,适用于任务的某个阶段存在多种方法选择的情形.

GOMS 模型使用上述四个组件描述了完成

常规任务所需要的技巧. GOMS 的潜在应用是相当广泛的. 本质上, GOMS 对一个任务的分析描述了人为了成功完成任务所必需的结构化知识. 在此基础上, 知道了操作的顺序后, 就可能对特定任务的执行时间做出定量的预测. 其他分析, 如错误的预测、功能性覆盖、学习时间等也可能由 GOMS 完成. 从 Card 提出原始的 GOMS 形式后, 出现了许多不同形式的 GOMS 分析方法, 每一个都有些小的改动^[6].

GOMS 模型是迄今为止最成功的人-计算机交互模型. 但是, 如果把 GOMS 模型用于虚拟现实环境, 那么该模型有以下不足之处.

GOMS 模型中, 用户的操作必须是精确的, 而虚拟现实环境中用户的操作大多是非精确的. 人类在日常生活中习惯于并大量使用非精确的信息交流, 人类语言本身就具有高度模糊性. 允许使用模糊的表达手段可以避免不必要的认识负荷, 有利于提高交互活动的自然性和高效性.

GOMS 模型对交互界面可用性评价的主要指标是时间. 但虚拟现实环境中, 交互的速度显然不是唯一的指标. 本研究认为交互有效性应该是评价虚拟现实界面更好的一个指标.

GOMS 模型中其交互界面和实际应用是相分离的. 而虚拟现实界面则和实际应用紧密相连, 任何一个虚拟现实界面就是一个具体的实际应用.

因此, 本研究扩展了 GOMS 模型, 模拟用户在虚拟现实环境中整合多通道信息, 通过系列操作完成期望目标的过程. 扩展后的模型的缩写为 eGOMS, 其中, G 代表 Goals, 表示用户期望达到的目标以及对目标的认知, 对目标的认知决定了用户将要采取的操作; O 代表 Operations, 表示用户为达到目标所完成的操作; M 代表 Modals, 表示交互过程中涉及的多通道信息; S 代表 Synthesis, 指特定情景下对多通道信息的整合. 由于虚拟现实环境下用户操作、动作的不精确性, 计算机必须依靠特定情景的约束才能理解用户意图. 图 2 是扩展 GOMS 模型的一个例子. 用户目标是绘制直线和圆相切. 用户通过手势绘制分离的直线和圆, 而通过声音发出“相切”命令, 而系统整合手势和声音通道完成直线和圆相切的目标.

对虚拟现实界面可用性的评价, 可用评估函数 $f(A, C, E)$ 来表示, 其中, A 表示注意 (Attention), 交互所需注意资源越少越好; C 表示认知负荷 (Cognitive load), 交互的认知负荷越低越好; E 表示交互有效性 (Effectiveness of interaction),

交互有效性越高越好, 而交互是否有效主要体现在交互的自然性、直接性和操作速度上. 交互越自然、越直接、速度越快, 那么交互有效性就越高.

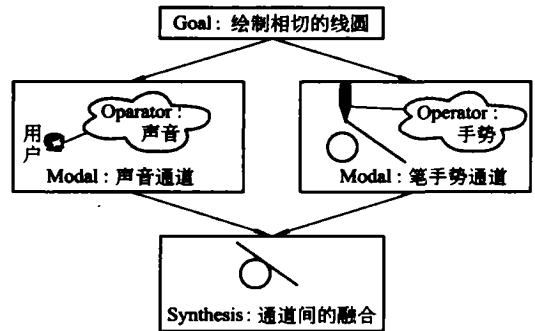


图 2 一个扩展 GOMS 的例子

3 基于 eGOMS 模型的 PIBG 交互范式评估

笔式输入由于符合人们千百年来形成的自然习惯而成为多通道用户界面研究中的热点. 从社会科学、认知科学的角度来看, 以笔为交互设备符合人的认知习惯, 承应了社会文化的氛围, 是一种自然高效的交互方式; 从技术的角度来看, 硬件技术已经成熟, 各种各样的笔交互设备不断出现; 软件技术的发展也促进了各研究领域的交融, 带动了相关学科的发展. 此外人们对纸笔还有特别的偏好, 这是因为其便携性、使用的手眼协调性 (视觉与手动的整合) 和直觉性 (草图的使用、思维连续性).

PIBG 是一种 Post-WIMP 交互范式^[1]. P 代表 Physical objects, 包括 Paper 和 Frame 等. 它的作用与 Windows 相当, 提供一个人机交互的基础; 但它更接近于人们的现实生活场景; IB 代表 Icons 和 Buttons. 其作用相当于 Icons 和部分 Menus 的功能. 一方面它们不可能像 Menus 一样提供大量的交互命令, 但却可以给用户很好的 Affordance 支持, 从而降低人机交互过程中用户的注意力和认知负荷; G 代表 Gestures. 一方面它取代了 Pointer 的功能, 另一方面它还完成了 Menus 部分的功能. 用户在发出命令时, 不必花费精力从菜单中寻找命令. PIBG 的心理学基础主要体现在以下四个方面.

a. PIBG 交互范式占用较少的工作记忆. 认知负荷是指在执行某种作业的过程中, 因作业特性所需要的认知能量或认知资源. 工作记忆是人类认知活动的重要资源. 工作记忆容量有限, 其中视觉工作记忆只能记忆 3~4 个特征或 3~4 个物体. PIBG 界面范式符合人类“工作记忆”的特点.

b. PIBG 交互范式占用用户较少的注意资源. 从注意的心理机制上看, 人可以灵活地分配注意资源去完成各种任务, 甚至同时做多件事情, 但前提是所要求的资源和容量不超过所能提供的资源和容量. 在 PIBG 交互范式的人机交互中, 用户的兴趣可以主要集中在其任务上, PIBG 交互范式占用很少的注意资源, 动作更易于形成自动化加工. 传统 WIMP 界面要求人迎合计算机的人机交互要求来转换、组织自己的思想, 将数据事先进行结构化处理, 这违背了人处理和表达脑内信息的自然习惯, 因此, 人们在从事知识创造方面的工作, 特别是在捕捉、组织和提炼信息时, 仍然习惯于使用纸笔.

c. PIBG 交互范式易于用户建立心理模型. 用户是根据从长时记忆(概念模型)中提取出来的知识以及当前环境信息, 在工作记忆中构建物理系统的心理模型. 用户凭借心理模型与系统进行交互, 当界面不直接提供完成任务所必需的信息时更为如此, 因此设计者的概念模型和用户的心理模型应该匹配才能避免用户错误和误操作. 笔式输入由于符合人们千百年来形成的自然习惯而使用户容易建立关于交互系统的心理模型.

d. PIBG 交互范式提供用户更多的 Affordance 支持. Gibson 认为动作和感知是通过 Affordance 来连接的, Affordance 是现实世界的物体为动物或人类提供的动作可能性^[7]. 在人类的进化过程中感知器官能够直接从环境中获取 Affordance(直接感知)而不需要进一步的认知过程. PIBG 交互范式提供用户更多的 Affordance 支持, 从而使得人机交互变得更加自然、和谐.

从认知心理学的角度来看, PIBG 交互范式(包括纸和框等交互组件)更加自然, 可以减轻用户的认知负担. 从用户笔输入的交互来看, 基于手势的交互方式更加自然和高效. 笔式输入与传统的键盘鼠标界面各自的特点, 但自由交互是笔式界面的一大优势, 笔式界面的自由互动形式非常适合人们进行一些捕获、组织和提炼信息方面的工作.

参 考 文 献

- [1] Dai G. Pen-based user interface[A]. In: Shen W, Li T, eds. Proceedings of the Eighth International Conference on CSCW in Design[C]. Xiamen, 2004. Beijing: International Academic Publishers, 2004. 32—36
- [2] 汪成为, 高文, 王行仁. 灵境(虚拟现实)技术的理论、实现及应用[M]. 北京: 清华大学出版社, 1996.
- [3] Tanriverdi V, Jacob R J K. VRID: a design model and methodology for developing virtual reality interfaces[A]. Proceedings of the ACM symposium on Virtual reality software and technology[C]. Banff, 2001. New York: ACM Press, 2001. 175—182
- [4] Baddeley A D. Working memory: looking back and looking forward[J]. Nature Reviews Neuroscience, 2003, 4(10): 1 829—1 839
- [5] Card S K, Moran T P, Newell A. The psychology of human-computer interaction[M]. Hillsdale: Lawrence Erlbaum Associates, 1983.
- [6] Carrol J M. HCI models, theories, and frameworks, towards a multidisciplinary science[M]. San Francisco: Morgan Kaufmann Publishers, 2003.
- [7] Gibson J J. The ecological approach to visual perception[M]. Boston: Houghton Mifflin, 1979.