

# 普适计算的人机交互框架研究

岳玮宁 董士海 王 悦 汪国平 王 衡 陈文广

(北京大学计算机科学技术系人机交互与多媒体研究室 北京 100871)

**摘 要** 普适计算通过向用户提供透明的计算服务使其注意力回归到任务本身. 它正在逐步发展成主流的计算模式. 人机交互的效率和自然性是影响其发展的关键因素之一. 文章从认知心理学的角度分析了用户在普适计算环境中的认知特征, 论证了人类有限的认知是普适人机交互中的瓶颈, 计算系统的交互模式必须适应用户的认知特征. 据此提出了普适计算模式的 Multimodal + Context-sensitive 交互框架, 对上下文信息和用户多通道主动输入信息进行无缝融合. 作者将其在移动导游导航系统 TGH 中加以实现. 用户评测结果初步证明该框架对于提高普适计算人机交互效率和自然性具有良好的效果.

**关键词** 普适计算; 认知特征; 上下文感知; 多通道交互  
**中图法分类号** TP391

## Study on Human-Computer Interaction Framework of Pervasive Computing

YUE Wei-Ning DONG Shi-Hai WANG Yue WANG Guo-Ping WANG Heng CHEN Wen-Guang

(Laboratory of HCI & Multimedia, Department of Computer Science and Technology, Peking University, Beijing 100871)

**Abstract** Pervasive/ Ubiquitous computing is emerging as an exciting new computing paradigm with the objective to provide computations everywhere at any time. It poses great challenge for current human computer interaction model. In this paper, the traditional user-driven interaction framework, in which human computer interaction exclusively depends on user's explicit input, is theoretically proved to be unsuitable for pervasive computing based on the analysis of characteristics of this new paradigm not only from the level of computing technology, but also user's cognitive behavior and psychology. Attention allocation of users essentially requires the computing systems to be more active, and the modality conflict needs more natural and flexible facilities. An interaction framework, in which the implicit knowledge extracted from context is seamlessly integrated with the explicit user input from multiple modalities, is then proposed to meet such requirements. It can be also regarded as an interaction design model for pervasive applications. Authors implement this framework in a mobile guide system TGH and evaluate its performance. The results show that this framework improves the naturalness and efficiency of interaction in pervasive computing.

**Key words** pervasive / ubiquitous computing; cognitive characteristics; context awareness; multi-modal interaction

收稿日期: 2003-10-08; 修改稿收到日期: 2004-09-28. 本课题得到国家自然科学基金重点项目(60033020)资助. 岳玮宁, 男, 1980 年生, 博士研究生, 研究方向为人机交互、模式识别、CAD. E-mail: ywn@graphics.pku.edu.cn. 董士海, 男, 1939 年生, 教授, 博士生导师, 研究方向为计算机软件、图形学、人机交互等. 王 悦, 女, 1979 年生, 硕士, 研究方向为人机交互与计算机图形学. 汪国平, 男, 1964 年生, 博士, 教授, 博士生导师, 研究方向为计算机图形学、CAD、人机交互等. 王 衡, 女, 1960 年生, 博士, 副教授, 研究方向为人机交互与图像处理. 陈文广, 男, 1964 年生, 博士研究生, 研究方向为人机交互.

## 1 引 言

计算模式在 20 世纪 80 年代经历了从主机计算(mainframe computing)到桌面计算(desktop computing)的革新,极大地推动了计算机技术和产业的发展.但是桌面计算的交互模式同时也给用户造成了诸多限制,使计算机仍然难以真正融入人们的工作和生活.这是当前计算机产业和应用领域进一步发展的主要障碍<sup>[1]</sup>.因此,计算机技术进一步发展迫切需要全新的计算模式,这就是普适计算<sup>[2]</sup>(pervasive / ubiquitous computing).它的目标是使计算设备和技术“消失”在用户日常生活和工作任务的背景当中,保证用户在得到计算服务的同时无需觉察计算机的存在和为此而分心,从而使其注意力回归到要完成的任务本身.以自然、和谐的人机交互环境为主要特征的普适计算是计算模式发展的大方向,具有广阔的发展前景.

计算机的“不可见”体现在技术上是计算系统的高嵌入性和高移动性<sup>[3]</sup>,硬件和网络环境是其基础.由于摩尔定律的存在,计算机硬件和网络在近年来都取得了迅猛的发展,而且这种快速的发展还将保持相当长的一段时间,可以认为建立普适计算环境的硬件支撑条件已基本成熟.但与此同时,人机交互的模式却绝大部分仍然停留在桌面计算的 WIMP(Window, Icon, Menu, Pointing device)范式下,繁琐而低效的交互过程严重地增加了用户的认知负担,干扰了其正常任务,从而影响用户对计算机应用系统的整体评价.因此,人机交互是普适计算中的一个关键问题.不断改善人机交互的自然性和效率,对于普适计算的发展具有重要意义.

作者认为,现阶段的普适计算人机交互主要受两方面因素的制约:

第一,新型计算设备的制约.一方面,小型化设备的使用日益广泛,其有限的交互空间和工具给交互过程带来了阻碍;另一方面,以智能家居、智能教室、智能房间为代表的大型嵌入式集成环境在很多领域内体现出了独特的优势.在这些智能环境中,用户在执行任务的过程中往往不断在各种不同设备之间切换,并且需要随时根据设备和任务特点变换交互手段.无论是小型化个人计算设备还是嵌入式集成空间,都对传统的交互工具和手段(如鼠标、键盘等)提出了挑战.

第二,用户认知特征的制约.设备的限制不是普

适交互问题的根本原因.如前文所述,计算机的不可见是普适计算的重要特点,而这种不可见并不都是计算机技术发展的直接结果,在很大程度上是用户心理作用的结果<sup>[1]</sup>.在普适计算模式中,用户在交互过程中的心理活动和认知特征与传统计算存在着很大差异,用户与周围环境和任务的关系更加紧密,人机闭环以外的诸多因素频繁且明显地影响着交互过程,这是导致普适计算人机交互诸多问题的深层次原因.

因此,要解决普适人机交互的问题,不但需要从计算机技术的角度入手,还必须从用户心理的角度进行讨论,通过适当的交互框架合理利用和整合交互技术,使交互方式适应用户的认知心理特征.基于用户认知心理特征的人机交互框架正是本文重点研究的内容.

## 2 相关工作

普适计算人机交互问题近年来受到了国内外的普遍关注.国际上,很多研究重点放在提高移动设备的交互能力上,从而使用户可以在任何时间、任何地点享受普适计算服务.美国斯坦福大学、CMU 和 Xerox 等研究机构始终关注 Web 信息在小屏幕设备上的呈现<sup>[4,5]</sup>;斯坦福大学和华盛顿大学则在提高普适计算环境的信息输入效率上也进行了不同的探索<sup>[6,7]</sup>.另一方面,将计算空间与物理空间紧密结合的嵌入式集成空间也得到了广泛的重视,MIT 的 Oxygen 项目中的智能房间,CMU 的 Aura<sup>[8]</sup>,斯坦福大学的 Interactive Workspaces<sup>[9]</sup>,欧盟的 Disappearing Computer<sup>[10]</sup>等研究计划是其中的主要代表.在国内,清华大学人机交互与媒体集成研究所构建的智能化教室 Smart Classroom<sup>[11]</sup>从人机交互和系统支撑软件的角度对普适计算的理论和实践进行了深入讨论,取得了良好的效果.中国科学院软件研究所在办公环境中基于笔的自然交互方面进行了有益研究<sup>[12]</sup>,并且已经进入产品化阶段.北京大学则在改善手持移动设备交互方面进行了一些探索性的工作<sup>[13]</sup>.

综合分析这些研究,我们发现:从人机交互技术的角度看,多通道交互和上下文感知是其中的关键问题.多通道交互一直以来就受到广泛的关注,美国加州大学伯克利分校的研究指出<sup>[14]</sup>:通过多通道方式访问 Internet 是普适计算中的关键问题之一.2002 年,国际标准化组织 W3C 成立的“多通道交互”

活动小组(Multimodal Interaction Working Group),旨在开发支持普适计算设备多通道交互的通用协议标准.它不仅详细描述了多通道交互的整体结构、输入构件和输出构件,还通过若干实例从实际应用的角度直观地说明了多通道交互对普适计算交互的重要作用.上下文感知方面的研究也很多,上下文的参与能够极大地提升计算系统的主动性和智能性.后文将对它们展开详细讨论.

在计算技术发展的基础上,真正实现计算机的不可见还需要通过合理的框架将这些计算技术进行融合,而这种融合的基础就是用户的认知心理特征.正如 CMU 在 Aura 项目中的研究指出的一样,普适计算中最珍贵的资源不是计算机技术,而是人的认知资源<sup>[8]</sup>.用户在普适计算交互过程中区别于传统模式的认知特征是所有普适计算系统所应该遵循的.因此,从心理学的角度分析用户的认知心理特征是十分必要的,它是普适计算人机交互的理论基础.作者目前尚未见到这方面研究的报道.

在普适计算交互的现有研究基础上,本文工作的特点在于:(1)从心理学的角度分析用户在普适计算环境中的认知特征,从而为普适计算人机交互提供了科学依据;(2)扩展多通道交互的概念范畴,将上下文信息作为一种特殊的交互通道,与用户输入进行统一的形式化表示和信息整合,实现了显式交互(explicit interaction)与隐式交互(implicit interaction)的无缝融合,并将其体现在一个通用的交互框架之中.

### 3 普适计算的用户认知特征

#### 3.1 注意资源的分配与竞争

注意(attention)是指人的心理活动对一定对象的指向和集中,即心理以清晰而生动的形式对若干种似乎同时可能的对象的一种占有<sup>[15]</sup>.它是一般认知活动的起点,即只有当注意集中到某个物体上时,对其后续的认知过程才可能展开.计算机在任何计算模式中都要同其它事物竞争用户的注意资源,只不过在传统桌面计算中,计算机几乎可以独占用户注意资源,因此这个问题表现的不突出.而普适计算中用户注意资源的分配特征却给现有交互模式提出了巨大的挑战,主要表现在如下两方面:

(1)注意资源的有限性决定了频繁的交互会影响用户任务的执行.人类的注意资源是有限的,注意的过程就是意识的汇聚与集中过程,它意指离开某

些事物以便有效地处理其它事物.也就是说当人注意某一事物时,就必然选择了忽视其它事物,指向性和集中性是注意的本质特征<sup>[16]</sup>.在普适计算中,用户的注意力回归到任务本身,他们与周围事物之间的联系比桌面计算紧密得多.因此如果计算机系统频繁地要求用户与之发生显式的交互,就会分散甚至中断用户集中在其正在执行的任务上的注意力,从而破坏了“透明性”要求.

(2)注意资源的选择性决定了用户集中在计算机上的注意力会明显下降.与桌面计算独占用户注意资源不同,普适计算中的计算机与其它任务在对用户注意的竞争处于同等的地位,而计算机在这种竞争中往往处于下风.从注意的生理机制上看<sup>[17]</sup>,注意是有机体的一种定向反射,反射的起点是感受器官受到新异刺激.而计算机在物理上的消失无疑会降低对用户感官的刺激.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.卡里曼(Kahneman)在其经典注意资源分配模型<sup>[18]</sup>中给出了注意资源在多种任务之间的分配原则,他指出:人对感兴趣或突发的事物会投入更多的注意,持久的倾向则是熟悉的事物容易吸引注意;同时,人们会对执行各种任务所需的能量进行估计,需要能量越少的任务越容易吸引注意,如图 1 所示.

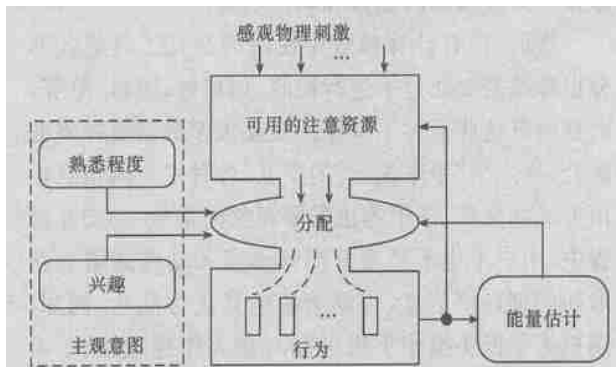


图 1 注意资源的分配模型

在普适计算环境中,用户的兴趣主要集中在其任务上,他们对这些任务的熟悉程度往往超过计算机,而且普通用户大多认为日常任务比操作计算机要简单.因此当用户集中精力完成某项任务时,其绝大部分注意资源都会集中在该任务而不是计算机上.注意力的下降直接导致了用户与计算机之间主

动交互的减少,而现有的交互模式中,用户主动交互是计算机执行操作最重要的驱动力.因此,如果交互模式不能变得更加主动,那么计算机在物理上的不可见就会导致其功能受到限制,甚至根本无法发挥作用,这就背离了普适计算的初衷.

我们通过数字智能教室中的一个例子进一步解释上述两个方面的影响.假设一个数字教室中的摄像头可以通过两个角度拍摄教室内的场景:一个拍摄教师的特写,另一个拍摄黑板上板书的特写.如果控制系统需要教师手动地选择拍摄角度,则可能会干扰其正常的教学活动和教学思路;反之,教师在讲授课程的过程中,很可能由于专注于教学本身而忽略了选择拍摄角度,从而使其功能无法发挥作用.

### 3.2 感官和动作通道竞争

无论交互模式变得多么“主动”,计算机与用户之间交互操作总是不可能完全避免的.在这些交互过程中,用户不但要进行注意资源的分配,还必须分配相应的感官和动作通道.而其它任务与计算机对用户通道的竞争是影响普适计算交互的又一个重要因素.心理学研究表明<sup>[17]</sup>,人必须接受并处理来自外界自然环境的信息并表现适应行为,且大部分环境信息是经过视觉传进大脑的,虽然人们可以通过其他感觉通道获取环境信息,但是视觉总是占有优势的;同时,在大脑加工这些信息之后,往往通过手进行相应的操作和反馈.人类总是相信眼睛和手是感知世界最可靠的感官和动作通道.

然而,现有计算模式中使用的交互工具绝大部分也都是需要通过手进行控制,如鼠标、键盘、笔等,而且所有这些通过手操作的交互大多数需要视觉的配合,称之为“手眼配合”的交互.各种主要的信息输出方式如文字、图形等也需要视觉的承接.在交互过程中,用户不得不经常分配其感官和动作通道去获取和处理环境信息,否则会影响其正常活动.例如,用户为了低头操作手机可能与他人相撞,这就不可避免地造成交互过程的不连续,从而严重地降低交互效率.因此,避免交互过程中的感官和动作通道竞争是提高交互效率和灵活性的关键问题.

## 4 普适计算交互框架

通过上小节分析可知,用户在普适计算环境中的认知特征决定了传统交互范式无法满足这种新型计算模式的需要.我们需要通过适当的技术和合理的交互框架来克服这些弊端.

### 4.1 普适计算的关键交互技术

#### (1) 感知和利用上下文

用户在普适计算中的注意分配特点决定了计算机应该更加主动,不应再将用户输入作为执行操作的唯一驱动力,它必须能够在用户不发出服务请求或是非精确交互的情况下智能地提供计算服务.在用户主动输入之外,交互上下文(context)是系统确定自身行为最直接、有效甚至是唯一的依据.上下文是指包括用户自身在内的计算环境中可能对交互过程产生影响的各种因素,其中蕴涵的丰富的环境信息能够拓宽人机之间的信息带宽,从而在交互过程发挥积极的辅助作用.它有助于系统对用户注意以外的任务进行智能化、自动化处理,缓解了注意干扰给交互过程带来的不良影响.普适计算环境中的上下文大致分为以下3类:

环境上下文:位置、速度、时间等环境因素.

设备上下文:网络带宽、屏幕大小等设备特性.

用户上下文:操作习惯、个人喜好、个性化需求等.

这3种上下文在交互过程中相互作用,对交互方式、用户界面进行自动调整,并对任务进行自动化处理而不干扰用户正常的认知任务.例如,移动导游中的位置信息就是一种重要的环境上下文,用户无需关注自己的具体位置,系统会自动检测并提供相应的服务.又如,对于一段待显示的文字,如果在掌上电脑或PDA上,我们可以采用普通的文字呈现方式输出,在手机、手表等屏幕过小的设备上则可以采用语音输出,设备上下文的这种参与有助于降低阅读疲劳.

上下文的感知过程是一个比较复杂的过程,对于环境和设备上下文,通常采用多级分层式结构模型进行感知,数据源是各种物理的或逻辑的传感器.用户上下文的获取是最困难的,它也是上下文交互的核心部分.用户上下文主要包括操作习惯、使用偏好、交互倾向、个性化需求等,对它们的感知本质上是一种用户学习的过程,学习的主要依据就是以往交互过程在时间上的积累,即上下文历史.对于较简单的应用系统,这种学习可以采用基于规则的算法实现;而对于那些复杂的应用系统,则需要利用人工智能中的相应学习算法.我们对上下文感知已进行过讨论<sup>[19]</sup>,因此本文将不在此深入展开.

#### (2) 基于语音和笔的多通道交互

多通道交互技术在普适计算中发挥着重要的作用.它允许用户通过各种自然的交互手段,如语音、手写等,与计算机系统进行交互,从而克服传统交互

工具缺失引起的问题.同时,多通道交互能够有效地扩大信息交换的带宽,从而达到提高交互效率的目的;并可发挥人机之间彼此不同的认知潜力,降低用户的认知负荷<sup>[20]</sup>.

在普适计算多通道交互中,语音是最重要的交互通道,它是普适计算系统用户界面的基础.这不仅因为语音交互最为自然,更因为它可以最大限度地解决通道竞争的问题.与包括笔在内的其它交互通道相比,语音独立性强,无需与手、眼配合即可执行交互任务,适合“眼忙手忙”条件下的应用,符合用户认知特征.当然,语音也存在本身的弱点和缺陷,例如目前语音识别技术还不是十分成熟;交流速度较慢;不易记录和修改编辑以及个人隐私等等问题.我们期望通过与笔的结合来解决这些问题,从而进一步改善普适计算人机交互.

笔也是一个具有极高自然性的交互通道,它既可以作为手写输入的载体,也可以在 WIMP 界面中

充当重要的指点设备.正是由于手写输入的自然高效以及 WIMP 界面存在的长期性,决定了笔在相当长的时间内还将是普适计算中的主要交互工具.更令人鼓舞的是,诸多研究表明笔和语音具有极强的多维互补性<sup>[21]</sup>,它们的结合使用能够极大地提高交互界面的易学性和灵活性,从而使用户可以根据环境和设备的特点灵活地选择最合适的交互方式.例如,用户在驾车的过程中,可以通过语音发送信息;而当下车来到办公场所后,出于不影响他人和隐私的考虑,可以改用手写或指点输入.作者认为基于语音和笔的多通道交互是普适计算人机交互的重要组成部分.

#### 4.2 普适计算的人机交互框架

充分考虑上下文参与的多通道交互符合普适计算环境中的用户认知特征,构成了普适计算人机交互框架的主体结构,如图 2 所示.我们将该框架的特点概括为 Multimodal + Context-sensitive.

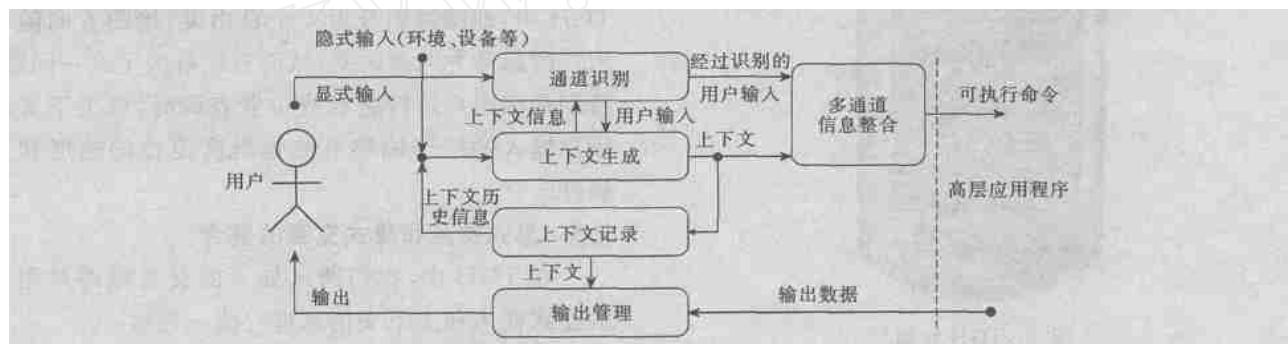


图 2 普适计算人机交互框架: Multimodal + Context-sensitive

在这个模型中,通道识别模块负责调用不同的子模块对用户通过不同通道输入的信息进行识别.上下文生成模块负责感知和生成上下文信息,上下文的来源主要包括 3 个方面:(1)用户主观以外的诸因素,包括自然环境因素和设备特征等;(2)经过识别的用户输入,它是用户意图和个性化因素最集中的体现,是用户上下文学习的主要依据;(3)上下文的历史记录,所有上下文形成后都被保存,从这些记录中往往可以提取规律性信息,挖掘出有意义的上下文.

需要特别注意的是,用户输入和上下文信息不是孤立存在的,而是一个有机的整体.从交互模式的角度考虑,多通道属于显式交互的范畴,其目的是使用户“易交互”;上下文则属于隐式交互,其目的是让用户“少交互”.只有通过两者的结合才能最终达到改善普适计算交互的目的,因此我们希望对两种信息进行无缝的融合.我们在信息整合阶段完成这种

融合.从广义的角度看,上下文也是一类特殊的交互通道,它是对用户主动输入信息的解释和补充.因此交互框架将上下文信息作为输入的一部分参与多元信息的整合,这样形成的系统命令既能体现用户的主观交互意图,又能充分考虑当时的客观因素.

为了对多源信息进行整合,需要对用户输入信息和上下文进行统一的规范化描述,并将其作为多通道信息整合模块的输入进行整合,整合算法可以参照文献[14]根据具体应用确定.在整合过程中需要注意,用户主动输入的优先级应当高于上下文信息,即如果二者之间存在冲突,应主要按照用户输入的语义形成命令.多元信息整合之后,系统形成了具有明确语义的交互命令,并根据这些命令调用相应的系统功能,而后将计算结果返回给用户输出信息管理模块.输出管理模块根据上下文历史中保存的上下文信息为输出选择最适合的媒体作为载体.

该框架并不是一个具体的体系结构,而是对交

互特征、关键构件和控制流程的一种抽象描述,如何实现可以根据具体需求而定.例如信息处理可以是分布式的,将某些耗费计算资源的操作放在服务器上,如某些输入通道的识别(如语音)、上下文的协同提取过程以及整合过程.

## 5 实 例

移动信息服务是普适计算的重要组成部分,为了验证上述交互框架的可用性,我们开发了智能导游系统 TGH(Tour Guide in Hand),采用 Microsoft Visual Embedded Tool 实现,运行于 Windows CE 兼容的移动设备上,具有无线网络通信功能以及基于 GPS 的地理信息查询功能.

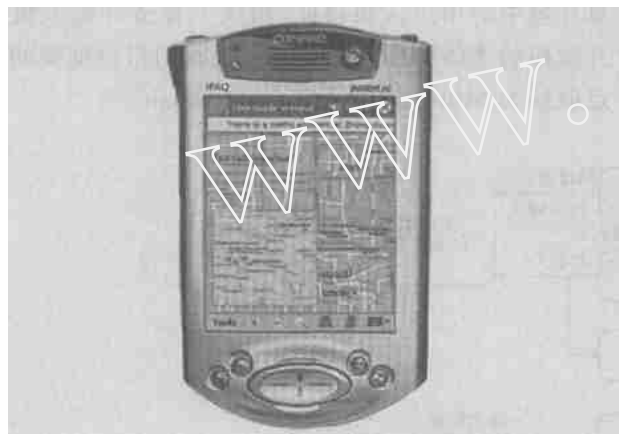


图 3 TGH 终端

### 5.1 TGH 中的多通道交互

TGH 支持用户通过语音、手写、指点、笔划手势以及它们的组合完成交互任务.其中语音的识别通过基于无线网的 C/S 结构实现,笔操作的识别在手持设备本地进行.用户可以通过不同的交互通道及它们的协同完成任务,具体采用哪个通道取决于命令类型和用户交互习惯.以查找到一个区域内的所有餐厅为例,用户可以通过笔在屏幕上指定一个区域,同时通过语音输入“餐厅”.又如查找位置 A 到位置 B 的最短距离,当用户确切知道 A、B 的名称时,可以通过语音输入“从 A 到 B 的路径”;若只知道其中一个(例如 A),则可以输入“从 A 到这里”,同时通过笔指点 B 的位置;若都不知道,可语音输入“从这里到这里”,并通过笔给出两点的位置.为了测试多通道交互的结果,我们邀请 36 位北京大学本科生使用 TGH 在北京大学校园内游览,对交互效果进行评估测试.测试结果表明<sup>[22]</sup>,多通道交互模

式较之单纯的笔和语音具有更高的交互效率.试验中我们发现,用户往往习惯使用笔完成那些较为简单的交互任务;而使用语音或两者的结合完成复杂任务,从而再次证明了两者的互补作用.

### 5.2 TGH 中的上下文交互

采用我们在文献[19]中提出的分层抽象模型,TGH 中实现了基于位置、方向、网络状况、游览路线、用户阅读浏览习惯等上下文信息的交互.这些上下文信息与用户输入共同作用,对系统的行为和界面进行自适应性的调整.例如,传统的 GIS 服务程序总是将正北作为地图的正方向,即无论用户向哪个方向运动,北方始终处于界面的上方.当用户向其它方向运动时,由于界面的方向表征(representation)与用户心理的方向表征之间出现了偏差,用户定位的速度和方向感会明显下降.例如用户向南行走时,物理世界中他左边的物体在界面上却显示在右边,用户在建立两种表征的映射关系时就会受到影响.TGH 中,地图被划分为 4 个自由度,地图方向随着用户的运动方向而改变,从而有效解决了这一问题.特别是在用户进行路线或位置查询时,该上下文和用户输入的结合能够有效地提高定位的速度和准确性.

### 5.3 显式交互和隐式交互的整合

在 TGH 中,我们通过如下的交互原语对用户

的显式输入和上下文信息进行统一表示:

原语 = 有效性 优先级 消息来源 时间 动作

描述 [ 输入参数 ]

有效性 = Yes| No

优先级 = 1| 2| 3| 4| 5

消息来源 = 用户输入 | 上下文

用户输入 = 语音 | 笔 | 其它通道输入

上下文 = 环境 | 设备 | 用户

时间 = time

动作描述 = message

输入参数 = 参数 1 [ 参数 2 参数 3 ...]

而后,我们采用基于任务制导的算法对这些信息进行整合并形成系统命令.举一实例,当用户试图查找当前位置到一个指定位置 A 的路线时,他可以通过语音输入“到这里”,同时通过笔指点目的地的位置.这时语音命令指明了任务的目的是查找路径,起点和终点是其两个参数;其第 1 个参数是用户的当前位置,由 GPS 接收并经过解析的数据得到;第 2 个参数是目的地,由笔通道的指点操作得到.

## 6 结 语

本文以普适计算人机交互为研究内容,从认知心理学和交互技术、框架的角度研究提高人机交互效率和自然性的方法.作者首先从认知心理学的角度分析了用户在普适计算环境中的认知特征,初步确立了普适计算人机交互的心理学基础.进而根据这些特征提出了通过基于语音和笔的多通道交互以及上下文感知提高交互效率的途径,并建立了普适计算环境人机交互的 Multimodal + Context-sensitive 框架,并在我们研究开发的移动导游导航系统 TGH 中得到实现.文章中提出的普适计算交互框架具有较强的通用性,有助于构建普适计算环境中的具有较高人机交互效率和自然性的交互式软件系统.

**致 谢** 衷心感谢北京师范大学心理学院舒华教授、牟书博士和北京大学李薇博士在认知心理学方面给予作者的重要指导和帮助!

## 参 考 文 献

- Xu Guang-You, Shi Yuan-Chun, Xie Wei-Kai. Pervasive computing. Chinese Journal of Computers, 2003, 26(9): 1042 ~ 1050 (in Chinese)  
(徐光祐, 史元春, 谢伟凯. 普适计算. 计算机学报, 2003, 26(9): 1042 ~ 1050)
- Weiser M.. Computer for the 21st century. Scientific American, 1991, 265(3): 94 ~ 104
- Lyytinen K. *et al.*. Issues and challenges in ubiquitous computing. Communications of ACM, 2000, 45(12): 62 ~ 65
- Buyukkokten O., Garcia-Molina H. *et al.*. Power browser: Efficient web browsing for PDAs. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '00), Hague, Netherlands, 2000, 430 ~ 437
- Partridge K., Chatterjee S. *et al.*. Tilttype: Accelerometer-supported text entry for very small devices. In: Proceedings of the 15th ACM Symposium on User Interface Software and Technology (UIST '02), Paris, France, 2002, 201 ~ 204
- Wobbrock J. O., Forlizzi J. *et al.*. WebThumb: Interaction techniques for small-screen browsers. In: Proceedings of the 15th ACM Symposium on User Interface Software and Technology (UIST '02), Paris, France, 2002, 205 ~ 208
- Guimbreti re F., Winograd T.. Flowmenu: Combining command, text, and data entry. In: Proceedings of the 13th ACM Symposium on User Interface Software and Technology (UIST '00), 2000, San Diego, USA, 213 ~ 216
- Carlan D., Siewiorek D. P., Smailagic A., Steenkiste P.. Project Aura: Toward distraction-free pervasive computing. IEEE Pervasive Computing, 2002, 1(2): 22 ~ 31
- Johanson B., Fox A., Winograd T.. The interactive workspaces project: Experiences with ubiquitous computing rooms. IEEE Pervasive Computing, 2002, 1(2): 67 ~ 75
- Salz P.. The disappearing computer. Time Europe, 2000, 155(8): 1 ~ 8
- Shi Yuan-Chun, Xie Wei-Kai, Xu Guang-You *et al.*. The smart classroom: Merging technologies for seamless tele-education. IEEE Pervasive Computing, 2003, 2(2): 47 ~ 55
- Li Yang, Guan Zhi-Wei, Dai Guo-Zhong *et al.*. A context-aware infrastructure for supporting applications with pen-based interaction. Journal of Computer Science and Technology, 2003, 18(3): 343 ~ 353
- Yue Wei-Ning, Dong Ai-Qin *et al.*. Multimodal Chinese characters input for handhelds. In: Proceedings of the 5th Asia Pacific Conference on Computer Human Interaction (APCHI 2002), Beijing, 2002, 652 ~ 658
- Ross S.J., Hill J.L. *et al.*. A compassable framework for secure multi-modal access to internet services from Post-PC devices. Mobile Networks and Applications, 2002, 7(5): 389 ~ 406
- Li Wei. Research on Cognitive Psychology. Hangzhou: Zhejiang People Press, 1998, 103 ~ 105 (in Chinese)  
(李 维. 认知心理学研究. 杭州: 浙江人民出版社, 1998, 103 ~ 115)
- Eysenck M. W.. Principles of Cognitive Psychology. Psychology Press, 1997
- Margaret W. M., Hugh J. F.. Sensation and Perception. Massachusetts: Allyn & Bacon, 1996
- Kahneman D.. Attention and Effort. New York: Prentice-Hall, 1973
- Yue Wei-Ning, Wang Yue *et al.*. Architecture of intelligent interaction systems based on context awareness. Journal of Computer-Aided Design and Computer Graphics, to appear (in Chinese)  
(岳玮宁, 王 悦等. 基于上下文感知的智能交互系统模型. 计算机辅助设计与图形学学报, 待发表)
- Dong Shi-Hai, Dai Guo-Zhong, Wang Jian. Human-Computer Interaction and Multimodal User Interface. Beijing: Science Press, 1999 (in Chinese)  
(董士海, 戴国忠, 王 坚. 人机交互与多通道用户界面. 北京: 科学出版社, 1999)
- Oviatt S. L.. Multimodal interface research: a science without borders. In: Proceedings of the 6th International Conference on Spoken Language Processing, Beijing, 2000, 3: 1 ~ 6
- Wang Yue, Yue Wei-Ning, Wang Heng, Dong Shi-Hai. Multimodal interaction in handheld mobile computing. Journal of Software, to appear (in Chinese)  
(王 悦, 岳玮宁, 王 衡, 董士海. 手持移动计算中的多通道交互. 软件学报, 待发表)



**YUE Wei-Ning**, born in 1980, Ph. D. candidate. His research interests include human-computer interaction, pattern recognition and multimedia.

**DONG Shi-Hai**, born in 1939, professor, Ph. D. supervisor. His research interests include computer software, computer graphics and human-computer interaction.

### Background

Interaction framework, as we know, plays foundational roles for the communications between human and computation. We notice that although the interaction tools, modalities, display modes and other techniques have greatly developed with the improvement of computing paradigm, the HCI(human computer interaction) framework does not change in essence. No matter in mainframe or desktop paradigm, user's explicit input exclusively drives the behaviors of computing system: the user tells the computer what s/he expects at a certain level of abstraction (e. g. by command-line, direct manipulation using a GUI, gesture, or speech input), and then the computers do it. Interaction process is simply an iterative user driven input-output circulation in human computer closed-loop.

In this paper, authors utilize the knowledge of cognitive psychology to theoretically analyze the users' cognitive characteristics in pervasive computing, which is the cognitive foundation of the pervasive interaction, and prove that current interaction model is unsuitable for this new paradigm. Because of the limitation and selectivity of human's attention allocation, computer systems in a pervasive environment should be more active

**WANG Yue**, born in 1979, M. S. . Her research interests include human-computer interaction and computer graphics.

**WANG Guo-Ping**, born in 1964, professor, Ph. D. supervisor. His research interests include computer graphics, CAD and human-computer interaction.

**WANG Heng**, born in 1960, associate professor. Her research interests include human-computer interaction and digital image processing.

**CHEN Wen-Guang**, born in 1964, Ph. D. candidate. His research interests include human-computer interaction.

and adaptive to provide computing services to the users without consuming too much of their attentions. The way to achieve this objective is context awareness. Due to the competitions between computer systems and physical objects on users' modalities, they prefer multimodal interface to ease the plight of users. Speech and pen are especially emphasized because their combination can help users to perform those hand-busy or eye-busy tasks.

Contextual information and multimodal information are not isolated but close associated with each other. The role of context is not only to improve the adaptation of systems, but also to give additional meaning and explain to user's input. The implicit knowledge through context and users' explicit interaction through multiple modalities ought to be seamlessly integrated in a general interaction framework. Such a framework is presented in the paper. It can be also regarded as a model to support universal interaction design for pervasive computing. Authors implement this framework in the mobile guide prototype TGH. User studies and performance tests show that its interaction efficiency and naturalness are higher those of traditional mode.