

在线双向适应的笔手势界面框架研究

徐礼爽^{1),2)} 田 丰¹⁾ 吕 菲¹⁾ 戴国忠¹⁾ 王宏安¹⁾

¹⁾(中国科学院软件研究所人机交互与智能信息处理实验室 北京 100190)

²⁾(中国科学院研究生院 北京 100039)

摘 要 提出一个在线双向适应的笔手势界面框架,该框架针对传统笔手势界面中静态手势识别器不能支持用户的个性化输入以及用户在笔手势界面中面临的手势记忆问题,提出了在线双向适应的策略:一方面系统能够适应用户(系统可以在线支持用户的个性化输入);另一方面用户可以学习系统(用户可以学习系统提供的某些笔手势).该框架包括 5 个部分:(1)双向适应笔手势输入解释模型;(2)双向适应笔手势输入解释流程;(3)上下文优先级定义;(4)纠错和模糊消解界面;(5)在线笔手势查询帮助系统.在该框架的指导下作者开发了一个原型系统并进行了对比实验评估.结果表明该框架在可用性上具有较大的优势.

关键词 自适应用户界面;笔式用户界面;笔手势交互;个性化交互;双向适应

中图法分类号 TP18 **DOI 号**: 10.3724/SP.J.1016.2010.01100

Research on On-Line Two-Way Adaptive Pen Gesture Interface Framework

XU Li-Shuang^{1),2)} TIAN Feng¹⁾ LV Fei¹⁾ DAI Guo-Zhong¹⁾ WANG Hong-An¹⁾

¹⁾(Intelligence Engineering Laboratory, Institute of Software, Chinese Academy of Sciences, Beijing 100190)

²⁾(Graduate University of Chinese Academy of Sciences, Beijing 100039)

Abstract This paper presents a framework for on-line two-way adaptive pen gestures based user interface. In order to solve the problem that the traditional stationary pen gestures recognition engine can't support users' personified gestures input habits and the problem that users maybe can't remember some gestures when using them, the authors propose an on-line two-way adaptive strategy, on the one hand, system can adapt to different users according to their different input habits(system supports on-line personified pen gestures input habits through mediate interface), on the other hand, users can learn the interpretation capability of the system(users can learn some system gestures through mediate interface or on-line pen gestures query and help system). The framework consists of five parts: (1) on-line two-way adaptive pen gestures input interpretation model; (2) on-line two-way adaptive pen gestures input interpretation flowchart; (3) contextual priority definitions; (4) context-based mediate interface; (5) on-line context-based pen gestures query and help system. The authors implement a prototype under the guidance of the framework; in such prototype users can input gestures to manipulate photos. A controlled comparative quantitative experiment based on the prototype is carried out. The result of the experiment validates the advantages of the framework in usability.

Keywords adaptive user interface; pen based user interface; pen gestures interaction; personified interaction; two-way adaption

收稿日期:2009-10-20;最终修改稿收到日期:2010-05-09. 本课题得到国家自然科学基金(60503054,60703078,U0735004)和国家“八六三”高技术研究发展计划项目基金(2007AA01Z158)资助. 徐礼爽,男,1982年生,博士研究生,主要研究方向为人机交互、笔式计算、自适应用户界面. E-mail: xulishuang@mails.gucas.ac.cn,lishuang03@ios.cn. 田 丰,男,1976年生,博士,副研究员,主要研究方向为人机交互技术、界面工具. 吕 菲,女,1984年生,硕士,主要研究方向为可用性工程、人机交互. 戴国忠,男,1944年生,研究员,博士生导师,主要研究领域为人机交互技术、计算机图形学. 王宏安,男,1963年生,博士,研究员,博士生导师,主要研究领域为人机交互技术、实时智能、软件工程.

1 引言

笔式用户界面 (pen-based user interfaces) 是 Post-WIMP 用户界面的主要形态之一, 笔式用户界面基于传统的纸笔隐喻, 为用户提供自然的交互特性, 用户可以通过手写笔来完成他们在诸多领域的想法和设计的快速勾画, 而不用关注于交互操作本身^[1]. 同时我们可以采用笔迹结构理解相关算法来识别和解释用户输入的手绘草图和符号^[2]. 笔式用户界面已经成功应用在思维捕捉和概念设计、文件审阅等众多系统中^[3-6]. 随着笔式用户界面研究的进一步发展及相关产品的进一步普及, 笔式用户界面还会应用于更多更广的领域.

笔手势交互是笔式用户界面的重要特征和交互方式. 笔手势是用笔画出的符号来调用计算机命令的技术. 用户使用笔手势来完成不同的任务, 如文字编辑、草图建模、UI 设计、3D 操作和导航等. 笔手势界面就是主要通过笔手势来进行交互的界面. 在本文的说明和描述中, 手势都狭义特指笔手势 (与虚拟现实系统中用手的姿势和行为来实现的手势交互无关). 笔手势交互具有诸多的优点, 具体表现在: 操作命令和操作对象可以用一个完整的笔画输入来完成, 这种交互方式速度快, 操作自然便捷; 笔手势由于其符号化的图形特征, 相对于文本形式的菜单命令更容易记忆; 笔手势可以应用于各种显示设备上, 从小的手持设备到面积大的白板和电子墙, 都可以方便地使用笔手势. 基于以上优势, 笔手势交互已经应用在诸多的笔式用户界面中. 笔手势具有以上交互优势的同时, 其应用现状仍存在问题, 表现在: ① 在用户采用笔手势进行交互时, 需要首先记忆这些笔手势, 而在设计手势集时, 设计出的笔手势间可能存在着相似性, 这种相似性可能会导致用户的记忆混淆, 因此需要好的机制来避免这种笔手势间的混淆, 保证笔手势的易学易用^[7]; ② 笔手势识别器拒绝识别或者识别错误时, 会导致用户的输入意图没有执行或执行错误, 如果没有提供纠错和模糊消解界面 (mediate interface) 来消除二义性或错误, 用户可能会对交互感到困惑^[8]; ③ 不同用户手势输入的习惯会存在差异, 可能会存在同一用户对手势集中两个不同手势的输入差异小于不同用户对手势集中某一手势的输入差异^[9], 而静态识别器不能解决这种由于用户间的输入差异而带来的识别问题. 分析笔手势存在的问题可以看出, 问题①是手势

设计时需要解决的问题, 问题②和③是用户在使用手势时需要解决的问题.

上述提到的笔手势在设计和使用时存在的问题, 已有不少相关研究. 针对设计时存在的问题, Long 提出并实现了一个手势设计工具^[10], 它能够设计易学易用的手势集, 例如它能够提示手势设计人员哪些手势间具有相似性并可能使用户产生混淆, 它能够提示哪些手势不容易学习和记忆, 哪些手势容易使系统产生识别错误等. 笔手势使用时存在的问题 (上述问题②和③) 和其他基于识别的界面 (如语音识别、手写识别等) 有相似之处, 针对识别器会出现的拒绝识别或识别错误等问题, 已有不少纠错和模糊消解的途径和方法. Cao 等提出了在线自适应的手势界面框架^[9], 该框架将在线自适应的手势识别器和交互上下文相结合, 并在实现中给出了具体的针对手势识别的纠错和模糊消解界面, 其具有以下优势: (1) 用户利用纠错和模糊消解界面动态提供手势样本以在线重训练手势识别器, 能快速地提高手势识别器的识别率; (2) 利用纠错和模糊消解界面收集用户反馈, 使得识别器能适应不同用户手写输入的个性化.

但目前笔手势的研究还存在着一些不足: Long 提出并实现的手势设计工具, 能为笔式界面设计良好的手势集. 但由于用户输入的个性化, 用户在使用笔手势时仍会出现上述的问题. Cao 提出的手势框架虽然提供了在线适应功能并且通过系统的评估实验说明了其框架的优势, 较好地解决了新手用户在使用笔手势时出现的手势识别和用户的个性化问题, 但是, Cao 提出的方式在以下两个方面有待扩展和改进: (1) 笔手势需要用户记忆, 用户在使用笔手势进行交互时缺少在线帮助; (2) Cao 提出的自适应的策略是系统完全被动地适应用户, 用户交互的过程是系统完全单向的适应用户的过程. 这种自适应的方式具有片面性, 因为在进行笔手势交互时, 当用户没有记住某些笔手势时, 用户不一定乐于自己定义所有的未被记住的笔手势, 如果用户可以通过查询帮助获得系统对应的手势, 并且用户觉得该手势易记易用时, 此时用户就会使用系统提供的笔手势. 同时, 如果用户觉得系统提供的手势不符合自己的输入习惯, 这时用户可以让系统适应用户的个性化输入. 相比于系统单向的适应用户的输入习惯, 这种系统和用户之间双向适应的过程对于新手用户会更加合理, 因为它趋向于人与人之间的自然交互过程, 最终达到系统和用户之间对手势输入的共识. 目

前这种双向适应的策略在某些场合已经得到应用, Pan^[11]等人针对多通道对话系统(multimodal conversation systems)中存在的普遍问题,即系统对多通道输入的解释能力的相对有限和不足,提出了使用双向适应的策略,使得在交互时用户和系统间能够动态地相互适应彼此的需求. 亦即一方面:系统在上下文的基础上帮助用户了解系统所能接受的处理能力,从而提高系统的可用性和健壮性;另一方面:系统通过在线学习用户多通道输入的新的表达方式,增强了整个对话系统对多通道输入的解释能力. 这种双向的适应过程最终达到系统和用户交互的共识,用户在学习到系统所提供的解释能力的同时,在此基础上系统学习用户新的表达方式,增强了系统对输入的解释能力. 初步的实验验证了这种双向适应的优势.

在对相关研究详细分析的基础上,本文提出了一个在线双向适应的笔手势界面框架,该框架独立于具体的笔手势界面. 框架包括系统对笔手势输入的完整的解释过程,在解释过程中遇到的手势识别错误时,采用了系统和用户间动态适应的双向适应策略,与此同时该框架提出了基于上下文的查询帮助功能,系统按优先级显示系统所支持手势集的图形化表示,用来辅助用户记忆.

本文第 2 节阐述相关工作;第 3 节详细描述框架;第 4 节介绍原型系统;第 5 节描述实验设计和评估结果;最后总结全文.

2 相关研究

在针对笔手势的研究中,很多研究试图通过开发笔手势的设计工具,以使笔式交互系统的设计者方便地创建和设计笔手势. Rubine 设计了 GRANDMA^[12]工具,这个工具可以让笔手势设计者通过多个样本来定义笔手势,使得设计者可以用较少的精力来设计基于笔手势的用户界面. Long 先后开发了 gdt^[13]和 quill^[10]笔手势设计工具,允许设计者输入和编辑训练样例,并对设计者输入的笔手势样本进行识别.

但是有研究者发现,尽管设计者能够很容易地利用工具创建和设计笔手势,用户在使用笔手势时,仍然存在很多问题. Long 对 PDA 用户使用笔手势的情况做了一次调查^[13],结果显示:用户认为笔手势是非常强大的、高效的、自然的;但用户希望有更多的笔手势得以使用,而且用户发现笔手势难以记

忆. 从研究者对用户使用 Tivoli 情况的调查中,我们看到笔手势的问题是“新手用户很难记住他们^[14]”. 我们认为,用户在使用笔手势时,不像菜单和按钮等形式的命令一样只需要用户再认(recognition),而是需要用户先记住(remember)笔手势,然后进行回忆(recall)并正确地画出来. 因此,笔手势的易学习性、易记忆性是笔手势可用性中非常重要的方面.

Long 在笔手势的易学性和易记忆性方面做了大量工作. 他首先假设,手势间的相似性会妨碍用户记忆手势. 因而,他做了一组实验,以研究什么因素导致用户认为两个笔手势是相似的^[15]. 对实验结果进行分析后,他提出了一个计算模型以预测笔手势之间的相似性. 并把这个计算模型应用于 Quill 工具之中. Quill 是用来帮助笔式用户界面的设计者创建并改进笔手势集的工具. Long 还做了一个实验研究,考察哪些因素影响了手势的易记忆性^[10],实验显示,图标性(iconicness 或形象性)是影响笔手势易学性和易记忆性的最重要的因素.

上述的研究都是在设计笔手势时考虑到的笔手势集的易学易用性. 并不能解决使用笔手势时出现的问题, Cao 针对笔手势使用时出现的问题,提出了在线的自适应笔手势框架,该框架有如下两个特点:(1)该框架将交互上下文和自适应笔手势识别器相结合,来提高笔手势的识别率.(2)该框架通过在线自适应来满足用户输入手势的个性化. 通过系统的评估实验说明,该在线自适应的笔手势框架能够较好地支持用户个性化输入. 但是 Cao 提出的自适应策略是系统完全被动地适应用户,用户交互的过程是系统完全单向地适应用户的过程. 而双向适应指的是用户和系统之间相互的适应过程,用户可以通过纠错和模糊消解界面及在线查询帮助系统来学习系统所能接受的输入,同时系统能够学习用户的输入习惯和交互方式,系统和用户双向适应的过程最终能达到系统和用户对界面操作的共识. 这种交互方式类比于人与人之间的交互方式,交互的形式自然高效. Pan^[11]等人将双向适应的策略应用到多通道对话系统中,增强了多通道对话系统对多通道输入的解释能力.

3 在线双向适应的笔手势界面框架

在本部分我们将详细介绍在线双向适应的笔手势界面框架,该框架描述了笔手势界面的交互过程,即系统对笔手势输入的解释,我们将解释定义为从

用户的输入到系统执行命令的过程. 整个的解释过程包含 3 个步骤, 依次为: (1) 用户笔手势输入; (2) 系统识别; (3) 执行命令或者调用纠错和模糊消解界面. 框架针对笔手势输入提出了相应的解释模型和解释流程. 同时, 在框架中我们对笔手势的上下文优先级进行了定义. 在该框架中还有一个相对独立的部分: 基于上下文的笔手势查询帮助系统, 用来辅助用户查询和记忆笔手势. 该框架适用于笔手势界面, 独立于具体的应用和实现. 整个框架包含 5 个部分, 分别是: (1) 双向适应的笔手势界面手势输入的解释模型; (2) 双向适应的笔手势界面手势输入的解释流程; (3) 笔手势的上下文优先级的定义; (4) 基于上下文优先级的纠错和模糊消解界面; (5) 基于上下文优先级的笔手势查询帮助系统.

3.1 双向适应的笔手势界面手势输入的解释模型

双向适应的笔手势界面手势输入的解释模型如图 1 所示: 集合 A 表示新手用户开始使用时用户的笔手势输入范围. 集合 B 表示新手用户使用前系统所能识别的笔手势输入范围, 可以看出, A 和 B 的交集 C 是新手用户刚使用时系统所能识别的用户输入. 随着使用过程的积累, 系统的识别范围和用户的输入范围随之更新变化, 共有以下几种情况:

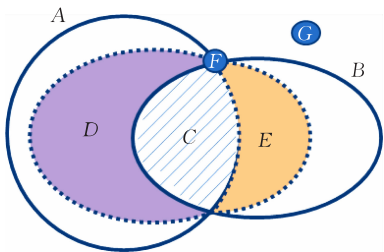


图 1 双向自适应笔手势界面手势输入的解释模型

(1) 当用户输入 $g \in C$ 时, 用户输入后的识别结果和用户的输入意图一致, 系统执行相应的任务. 如表 1(1) 所示.

(2) 当用户输入手势 $g \in D$ 范围时, 系统拒绝识别用户的输入, 而此时用户的输入意图是刚开始时 $G' \subset B-C-E$ 的手势集所对应的任务, 其中 G' 为对应任务的手势类的输入范围集合, 由于输入 $g \in D$ 在 B 之外, 系统拒绝识别输入的手势, 用户通过纠错和模糊消解界面, 将 g 作为样本添加到识别器所对应的手势类的样本集合中, 如果是首次, 则还需向识别器添加对应手势类的负类(我们将手势的负类定义为与该手势类别所对立的手势类, 我们用正整数来编号手势类别, 某手势负类的编号是该类手势编号的相反数), 并删除该类起始时所对应的样本并

将这些样本作为对应的负类的样本. 如表 1(2)“适应用户”列所示. 通过这种途径, 系统将用户输入的手势添加到相应的手势类的样本集合, 使得系统学习用户的输入习惯, 系统所能处理的 $B-C-E$ 范围对应的笔手势类将向 D 迁移, 最终达到 $B-C-E$ 到 D 的完整变迁.

(3) 当用户输入笔手势 $g \in A-C-D$ 时, 系统未能识别输入的手势, 用户通过错误模糊消解界面, 认为此时系统能够识别的 E 手势类集合的笔手势相比用户的手势输入有更好的易学易用性, 用户在这个过程中选择学习系统提供的笔手势, 如表 1(2)“学习系统”列所示. 用户输入的范围从 $A-C-D$ 向 E 迁移, 最终达到 $A-C-D$ 到 E 的完整变迁.

(4) 用户输入 g (图 1 没有显示这种情况), 系统识别错误时(用户意图输入为 G_x , 识别成 G_y), 这种情况可能为: 用户对 G_x, G_y 间手势记忆混淆; 识别器在识别时存在着误差, 导致的原因可能是设计手势集时 G_x 和 G_y 之间存在着相似性 ($G_x \cap G_y \neq \emptyset$). 撤销命令后, 用户调用纠错和模糊消解界面, 可能会出现以下两种情况: ① 用户发现自己混淆了手势, 将 G_y 所对应的手势记成了 G_x 所对应的手势, 用户通过反馈界面直接执行命令 G_x 所对应的命令, 并且记忆 G_x 所对应的手势, 这种情形对应表 1(3)“学习系统”列; ② 用户的意图输入为 G_x 所对应的手势, 系统误识别成 G_y , 此时用户会将 g 添加到 G_x 的样本集中, 将 g 添加到 G_y 的负样本中. 此种情形下系统学习用户的输入习惯, 最终系统会消除这种输入歧义. 如表 1(3)“适应用户”列.

(5) 刚开始用户输入 $g \in F$ 时, 用户当前的输入是非意图输入, 被识别成命令, 用户撤销相应的命令, g 作为非手势类样本(非手势类和手势类相似, 其代表该类为系统所能接受手势类, 但该类手势的输入不执行系统的任何命令), 如表 1(4) 所示.

(6) 用户输入 $g \in G$, 当前的输入是非意图输入, 没有识别成命令, 如表 1(5) 所示.

表 1 给出的双向适应的手势识别器适应策略, 和解释模型存在着对应. 适应策略基于 Rubine 手势识别算法, O 表示识别器拒绝识别或用户的无意图输入, 该算法可以在线为某类手势添加样本, 也可以根据需要在线添加某些手势的负类. 如果手势识别的结果为某一负类(编号为负数), 表明识别器拒绝将用户的输入识别成为手势. 采用其他手势识别算法, 其适应策略与表 1 的策略相似.

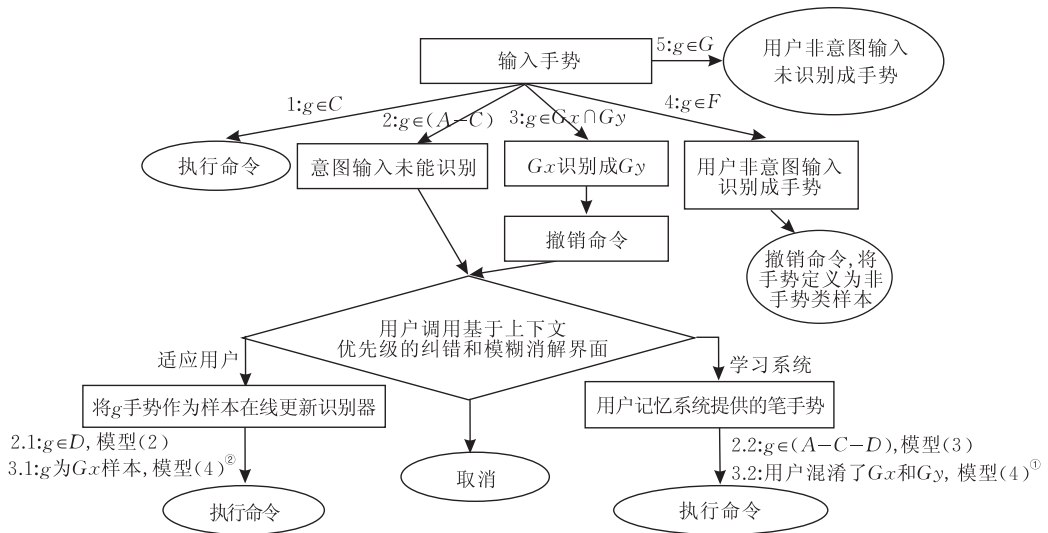
表 1 双向适应的手势识别器适应策略

情形	描述	用户意图	识别结果	学习系统	适应用户	取消
(1)	正确	G_x	G_x			
(2)	错误	G_x	O	用户记忆系统提供的 G_x 手势	将当前输入手势添加到 G_x 的样本集,如果是首次(添加手势类为 G_x 的负类,删除 G_x 的以前样本,并将 G_x 的以前样本添加到 G_x 的负类)	
(3)	错误	G_x	G_y	用户记忆系统提供的 G_x 手势(用户混淆了 G_x 和 G_y)	将当前输入手势添加到 G_x 的正样本集,如果是首次(分别添加手势类为 G_x 和 G_y 的负类,将 G_x 的现有样本设为 G_x 的负类样本,将 G_y 的现有样本设为 G_y 的负类样本,删除 G_x, G_y 类的现有样本)	
(4)	错误	O	G_x			将手势定义为非手势类样本
(5)	正确	O	O			

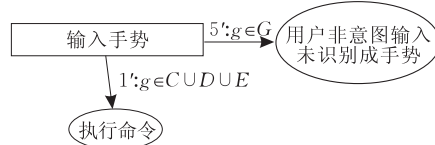
3.2 双向适应的笔手势界面的手势输入解释流程

如图 2 所示为系统对手势输入的解释流程,图 2(a)为新手用户开始使用时的解释流程图,我们

将输入分为 5 种情况 1,2,3,4,5. 每种情况对应的分支流程为该情况下手势输入的解释流程.



(a) 系统开始时手势输入的解释流程



(b) 系统输入稳定后手势输入的解释流程

图 2

此 5 种情况和上述模型存在着对应,即 1→模型(1),2→模型(2),(3),3→模型(4),4→模型(5),5→模型(6). 基于上下文优先级的纠错和模糊消解界面有 3 种反馈:“适应用户”;“学习系统”;“取消”. “适应用户”和“学习系统”的策略因流程而异. 流程 2→2.1 适应用户的策略为模型中(2)提出的策略;流程 2→2.2 学习系统的策略为模型中(3)提出的策略,流程 3→3.1 学习用户的策略为模型中(4)^②中提出来的策略,流程 3→3.2 学习系统的策略为模型中(4)^①提出的策略.

初始时 1,2,4,5 对应的值分别为 $g \in C, g \in A-C, g \in F$ 和 $g \in G$, 3 的输入为 $g \in G_x \cap G_y (G_x \cap G_y \neq \emptyset, x \neq y, 0 \leq x, y < m, m$ 为手势类的个数). 整个流程随用户的输入而进行迭代,迭代的过程中解释流程不变,而变量 1,2,3,4,5,2.1,2.2,3.1,3.2 的值随输入逐渐更新,如开始时,用户输入 1 为 $g \in C, 2$ 为 $g \in A-C$,随着用户使用经验的积累,1 的范围逐渐扩大,2 的范围会逐渐减小,直至最后 1 的输入范围为 $g \in (CUDUE), 2$ 的输入范围为 $g = \emptyset$. 如果开始时 3 为 $g \in G_x \cap G_y (G_x \cap$

$G_y \neq \emptyset$), 在用户和界面交互的积累过程中, $G_x \cap G_y$ 会不断减小, 最后为 \emptyset .

我们可以看出, 手势输入的解释模型静态地描述了用户和系统双向适应后用户的输入手势范围的变化, 而手势输入解释流程以流程图的方式动态描述了双向适应的过程以及过程中的变量发生的变化, 在用户和系统达成共识以后(用户和系统达成一致的输入范围 $g \in (CUDUE)$), 手势输入的解释流程如图 2(b) 所示, 相比于开始流程图, 分支 2、3 和 4 已经消除, 系统达到较高的手势识别率, 相比于开始使用时交互效率有显著提高.

3.3 基于上下文的优先级定义

用户和系统交互时蕴含着丰富的上下文信息, 如由交互历史构成的交互上下文, 由设备信息和环境信息构成的设备上下文和环境上下文等. 在交互上下文中, 由用户和系统在交互过程中组成的命令序列构成了用户和系统的交互历史, 在交互历史中蕴含着用户和系统的交互模式. 同时笔手势交互具有其特殊性, 我们将对笔手势的记忆特征进行建模. 笔手势的记忆特征模型用来表明用户对每一个手势的记忆特征. 同时, 我们用一个概率函数来计算输入手势与每一个手势类的相似度. 本文的上下文包括 3 个部分: 第 1 部分为交互上下文的建模, 即基于交互上下文的命令预测; 第 2 部分为笔手势记忆特征建模; 第 3 部分为手势的相似度函数.

3.3.1 基于交互上下文的命令预测

交互上下文由交互的历史序列组成, 交互历史动态增量地对系统执行的命令进行累积记录. 交互历史是对用户交互习惯的建模. 我们将每一个命令封装成为一个对象, 命令对象的种类由一个关键字来表示, 交互历史表示为执行的命令序列所对应的关键字序列, 每一次命令的执行会动态地增加这个关键字序列.

我们将基于交互上下文的命令预测定义为: 基于当前的交互上下文, 用户接下来执行某一个命令的概率. 其中上下文用 I_h (Interaction history) 来表示, 命令 $c_i \in C$ ($0 \leq i < |C|$) (C 为命令对应的关键字集合), $I_h \in S\{S = c_i S, S = c_j, c_i, c_j \in C\}$, 则在交互历史 I_h 下, 系统当前执行命令 c 的概率为 $P(c | I_h)$. 我们在此采用一种统计模型来计算这种概率, 该模型为部分匹配预测算法 (PPM), 具体的算法见文献 [16].

3.3.2 手势的记忆特征建模

在笔手势界面中, 用户在进行笔手势输入时需

要记住手势的图形化表示. 对于新手用户来说, 用户对系统提供的手势会有记忆上的误差, 容易忘记或者混淆某些手势. 我们提出, 在新手使用的过程中, 他们对每一类笔手势都有相应的记忆特征, 对手势的记忆随着交互过程的积累会逐渐加深, 在此我们用该手势使用时寻求帮助和通过纠错和模糊消解界面的反馈的次数之和与该手势使用的总次数的比来定义笔手势的记忆特征, 如公式所示, $M(g)$ 的值越大, 说明该手势越难以记忆.

$$Memory(g) = (Num_{help}(g) + Num_{mediator}(g)) / Num(g)_{execute} \quad (0 \leq Memory(g) \leq 1, g \in G).$$

3.3.3 手势相似度函数

当用户输入手势 g 时, 如果系统拒绝识别, 我们仍能计算其与系统每个手势类的匹配程度, 匹配程度高说明两个手势间的相似程度高, 在此我们使用 Rubine 算法中的匹配代价^[10]来计算手势与手势类的匹配程度, 或称相似度为 $Match(g, g_i)$ ($0 \leq i < |G|$), 因此输入手势与现有某手势类的相似度函数定义为

$$Similarity(g, g_i) = Match(g, g_i) / \sum_{j=0}^{|G|} Match(g, g_j) \quad (0 \leq i < |G|).$$

3.3.4 基于上下文的优先级

本框架两个部分需要上下文优先级的支持, 第 1 个是基于上下文优先级的纠错和模糊消解界面, 第 2 个是基于上下文优先级的在线笔手势查询帮助系统. 由于这两者有所区别, 纠错和模糊消解界面的上下文包括用户输入的笔手势信息, 而查询帮助系统没有此信息. 因此我们分别定义两种情况的优先级.

纠错和模糊消解界面的上下文优先级定义如下, 其中 g 为用户当前的手势输入:

$$Priority_{mediator}(g_i) = x \times Similarity(g, g_i) + y \times P(g_i | I_h) + z \times Memory(g_i) \quad (g_i \in G \text{ 且 } 0 \leq i < (|G|)).$$

笔手势查询帮助系统的上下文优先级定义如下:

$$Priority_{help}(g_i) = m \times p(g_i | I_h) + n \times Memory(g_i) \quad (g_i \in G \text{ 且 } 0 \leq i < (|G|)),$$

其中 $P(g_i | I_h) = P(c_i | I_h)$. x, y, z, m, n 为系数, 在此我们经验性地分别将它们赋值为 0.3, 0.2, 0.5, 0.3, 0.7. 直观上 $p(c_i | I_h)$ 为命令执行的可能性, 它的值对该手势是否需要帮助的作用相对不太明显, 我们将 $Memory(g_i)$ 所对应的项的系数设为比较大

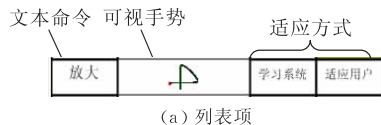
的值,因为该函数比较清楚地说明了某个手势需要帮助的可能性.

3.4 基于上下文优先级的纠错和模糊消解界面

当用户的输入意图没有得到执行时,用户通常需要通过纠错和模糊消解界面消除错误或歧义,手势界面中的纠错和模糊消解界面解决的是手势识别时存在的问题,该界面同时还要满足支持双向适应的需要.我们的纠错和模糊消解界面是一个列表,列表的每一项为系统支持的一种手势,如图 3(a)所示,每一项包括文本命令、可视手势、适应方式 3 个要素,其中适应方式为学习系统或适应用户两种,列表的顺序基于手势的上下文优先级从高到低进行排序.优先级的计算采用 $Priority_{mediator}(g_i)$ 计算公式.一个原型应用系统的纠错和模糊消解界面如图 3(b)所示,其显示了当前的用户输入,以及按照优先级顺序排序的手势列表,用户通过选择某一项的“学习系统”或“适应用户”来支持双向适应策略,同时如果用户想放弃这次输入,用户只需选择“取消”按钮.

3.5 基于上下文优先级的在线笔手势查询帮助系统

和菜单或工具条相比,笔手势是不可见的,当用户使用笔手势界面时可能没记住输入命令对应的笔手势,此时用户可以通过基于上下文的笔手势查询帮助系统来寻求帮助,我们的笔手势查询帮



(a) 列表项

当前输入为			
放大		学习系统	适应用户
重做		学习系统	适应用户
上翻		学习系统	适应用户
批注		学习系统	适应用户
查询		学习系统	适应用户
撤销		学习系统	适应用户
还原		学习系统	适应用户
下翻		学习系统	适应用户
放入		学习系统	适应用户

(b) 基于上下文优先级的纠错和模糊消解界面

图 3

助系统按照上下文的优先级从高到低列出所有手势对应的文字命令和手势的图示.上下文的计算采用 $Priority_{help}(g_i)$ 的计算公式,较高的优先级说明在该状态下,用户寻求该手势帮助的可能性高.一个原型系统的上下文优先级的笔手势查询帮助系统如图 4(b)所示.



(a) 原型系统界面

(b) 查询帮助系统界面

图 4

4 原型系统

我们在此框架下,开发了一个原型系统.该系统以框架为指导,实现了通过手势输入来管理图片的功能.该原型系统功能丰富,通过该系统的使用,用户能够在较短的时间内全面了解本文提出的框架,从而也方便了我们通过实验来对框架进行评估.该

原型系统具有的功能有:用户可以将图片从左边的图片架上移入右边的收藏集里,反之,从收藏集中将某图片放入图片架中;用户可以上下翻页导航浏览所有图片;放大还原某一张图片;给某张图片添加手写注释和查询手写注释等.所有这些任务都是通过二维的单笔手势操作来完成,操作对象由手势的几何中心决定.图 4(a)显示了该原型系统的一个截图.在该图中一个放大命令的手势作用在图片架右

下方的一张图片上. 根据任务我们定义了 9 个手势, 从图 3(b) 或图 4(b) 中我们可以看出. 用户通过双击笔尖调用查询帮助界面.

5 实验评估

5.1 实验描述

我们将本文提出的在线双向适应的笔手势框架和 Cao 提出的在线自适应手势框架以及无自适应功能的笔手势界面进行了实验对比. 我们的 3 个测试程序都是用 C++ 来实现的. 实验的手势识别器基于已有的 Rubin 手势识别算法. 在实验的过程中, 我们召集了 12 名大学生作为被试, 其中 7 名男生, 5 名女生, 他们先前都没有笔式输入的经验, 且都是右手输入习惯.

在每一个被试实验的过程中, 我们的图片管理程序呈现给用户不同的图片组, 每个组都是固定的 24 张图片, 每名用户需要通过手势输入完成如表 2 所示的任务.

表 2 原型任务描述

任务序号	任务描述
01	将左边图片库中的某一张图片放到右边的图片集中
02	将右边的图片集中的某一张图片放到左边的图片库里
03	对左边的某一张图片进行自由手写注释
04	查看已注释图片的注释信息
05	放大然后还原左边的某一张图片(两个子任务: 放大和还原)
06	上翻, 使左边的图片库显示的图片整体向上滚屏一行
07	下翻, 使左边的图片库显示的图片整体向下滚屏一行

在实验的开始阶段, 先向被试介绍原型系统, 被试同时阅读书面的任务描述. 在阅读阶段完成后, 向被试展示手势集, 同样用书面的形式来解释所有的手势, 由于我们的目的是研究新手用户在应用时的效率和效果, 因而没有让用户来练习手势的使用. 用户按照自己的进度来完成这些任务, 在实验的各个阶段之间, 给予用户足够的休息时间, 对用户输入的每个笔手势, 我们记录下识别结果的正确性以及手势完成所需时间.

实验采用了被试间设计. 自变量是: 在线适应的方式, 自变量有 3 个水平, 分别对应着: (1) 本文提出的在线双向适应的笔手势界面框架(G2); (2) Cao 提出的在线自适应笔手势界面框架(G1); (3) 无自适应功能的笔手势界面(G0). 我们按照自变量的 3 个水平将被试人员随机分成 3 组. 分别对应如下.

第 1 组: G2; 第 2 组: G1; 第 3 组: G0;

每一组被试都需要完成 3 个阶段的实验, 阶段

间允许间歇. 在每个阶段, 每个任务分别完成 4 次, 共 $4 \times 8 = 32$ 次任务执行, 任务的顺序按约束的规则随机产生. 每名被试在 3 个阶段完成规定的 $32 \times 3 = 96$ 次任务. 总体上说, 我们的实验如下:

3 组(自变量的 3 个水平) \times 4 个被试/组 \times 3 个测试阶段/每个被试 = 36 个测试阶段.

5.2 实验结果

我们将识别率定义为在一个阶段测试中, 被试正确输入手势的个数占被试输入手势总的个数的百分比.

使用 3 种不同自变量水平的被试组完成任务的识别率和操作时间的均值如表 3 所示, 使用 One-Way ANOVA 分析得出, 双向适应的框架比 Cao 提出的自适应框架在识别率上有显著提高 ($p < 0.0001$), 与无自适应功能的笔手势界面相比也有显著提高 ($p < 0.001$), 另外, Cao 提出的自适应框架识别率也显著高于无自适应功能的笔手势界面 ($p = 0.045$), 这与 Cao 的实验结果^[7]是一致的. 同时比较 G0、G1 和 G2 的标准差, G2 也更稳定.

另外, 在操作时间方面, 双向适应的框架显著快于 Cao 的自适应框架及无自适应功能的笔手势界面 ($p = 0.045$, $p < 0.0001$), Cao 的自适应框架也显著快于无自适应功能的笔手势界面 ($p < 0.0001$), 这与 Cao 的实验结果^[7]也是一致的.

表 3 识别率和操作时间

	识别率/%	识别率的标准差	平均时间/ms	时间的标准差
G2	95.27	0.04542	821.06	1295.280
G1	78.63	0.10162	1281.86	1147.960
G0	77.30	0.04722	2255.03	4713.777

图 5 显示了 3 种自变量水平下用户完成 3 个 session 所需的时间. G2 在 3 个 session 的操作时间都少于其他两种条件. 从整体趋势上, 我们可以看到随着用户操作的熟练, 用户的操作速度随熟练的程度越来越快.

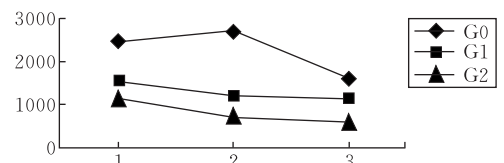


图 5 用户在 3 种自变量水平下不同 session 所花时间 (ms)

6 结论及未来工作

本文提出了双向适应的笔手势界面框架, 同时

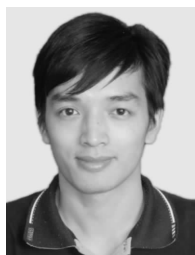
实验说明了该框架具有较好的可用性. 我们在实验的过程中发现, 用户在记不住某些命令对应的手势时, 有时会采纳系统提供的笔手势, 有时采用自己定义的笔手势, 用户没有记住或输入错误某种笔手势时并不总是乐于定义个性化的笔手势, 而有时会采用系统提供的笔手势. 这也说明了双向适应过程的必要性.

在实验中我们调查了用户的反馈意见, 有些用户觉得纠错和模糊消解界面交互起来比较复杂, 影响了用户的使用. 怎样提供便捷方便的纠错和模糊消解界面将是我们未来的工作之一. 为了实验控制的需要, 本实验将整个过程分为三个阶段, 每一阶段内每个任务执行 4 次, 但任务间的顺序是随机生成的, 实验过程中的交互上下文可能与在实际应用中的交互上下文有所区别. 更加系统深入地评估实验也是未来的工作之一.

参 考 文 献

- [1] Hong Jason I, Landay James A. SATIN: A toolkit for informal ink-based applications//Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology. San Diego, California, United States, 2000; 63-72
- [2] Alvarado C. Multi-domain sketch understanding. Massachusetts Institute of Technology, Massachusetts, United States, 2004
- [3] Landay J A, Myers B A. Interactive sketching for the early stages of user interface design//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Denver, Colorado, United States, 1995; 43-50
- [4] Hong J A, Landay J A. SATIN: A toolkit for informal ink-based applications//Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology. San Diego, California, United States, 2000; 63-72
- [5] Lin James. DENIM: Finding a tighter fit between tools and practice for Web site//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. The Hague, The Netherlands, 2000; 510-517
- [6] Tian Feng, Qin Yan-Yan, Wang Xiao-Chun, Ao Xiang, Wang Hong-An, Dai Guo-Zhong. Analysis and design on PIBG toolkit: A pen-based user interface toolkit. Chinese

- Journal of Computers, 2005, 28(6): 1036-1042
- [7] Long Jr A C, Landay J A, Rowe L A. Implications for a gesture design tool//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Pittsburgh, Pennsylvania, United States, 1999; 40-47
- [8] Mankoff J, Hudson S, Abowd G. Interaction techniques for ambiguity resolution in recognition based interfaces//Proceedings of the 13th Annual ACM Symposium on User Interface Software and Technology. San Diego, California, United States, 2000; 11-20
- [9] Cao Xiang, Ravin Balakrishnan. Evaluation of an on-line adaptive gesture interface with command prediction//Proceedings of the 2005 Conference on Graphics Interface Victoria. British Columbia, 2005. School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada; Canadian Human-Computer Communications Society, 2005; 187-194
- [10] Long Jr A C. Quill: A gesture design tool for pen-based user interfaces. University of California, Berkeley, CA, USA, 2001
- [11] Pan Shimei, Shen Siwei, Zhou Michelle X, Houck Keith. Two-way adaptation for robust input interpretation in practical multimodal conversation systems//Proceedings of the 10th International Conference on Intelligent User Interfaces. San Diego, California, USA, 2005; 35-42
- [12] Rubine D. Specifying gestures by example//Proceedings of the 18th Annual Conference on Computer Graphics and Interactive Techniques. New York, USA; ACM Press, 1991; 329-337
- [13] Long Jr A C, Landay J A, Rowe L A. PDA and gesture use in practice: Insights for designers of pen-based user interfaces. University of California at Berkeley, Berkeley, CA, USA; Technical Report UCB//CSD-97-976, 1997
- [14] Pedersen E R, McCall K, Moran T P, Halasz F G. Tivoli: An electronic whiteboard for informal workgroup meetings//Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems. Amsterdam, The Netherlands, 1993; 391-398
- [15] Long Jr A C, Landay J A, Rowe L A, Michiels J. Visual similarity of pen Gestures//Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. The Hague, The Netherlands, 2000; 360-367
- [16] Cleary J, Witten L H. Data compression using adaptive coding and partial string matching. IEEE Transactions on Communications, 1984, 32; 396-402



XU Li-Shuang, born in 1982, Ph. D. candidate. His research interests include pen based interaction, adaptive user interface.

TIAN Feng, born in 1976, Ph. D., associate research fellow. His current research interests include pen based interaction, human computer interaction.

LV Fei, born in 1984, M. S.. Her research interests include usability engineering, human computer interaction.

DAI Guo-Zhong, born in 1944, Ph. D., professor, Ph. D. supervisor. His current research interests include

computer graphics, human computer interaction.

WANG Hong-An, born in 1963, professor, Ph. D. su-

Background

Pen based user interface is a main kind of Post-WIMP user interfaces and many applications have adopted pen based interaction due to its free style interaction. Pen gestures interaction is very important in pen based user interface. There are many researches on the topic of pen gestures.

Some researches are about the algorithms to classify gestures such as rubine algorithm and a one dollar recognizer which both can effectively classify pen gestures. There are other researches which concentrate on how to design pen gestures set to facilitate the usability of pen gestures. Allan Chris Long has done lots of this kind of works and has proposed tools to facilitate the design of pen gestures set for pen based user interface.

But there are also some problems when we use pen gestures, they are: (1) different users have different gesture input habits, but the above algorithms don't support dynamic adaptive function to accommodate users' personified habits. (2) Users maybe don't know or forget the specific gestures, but systems don't offer on-line query and help function to assist users to use gestures.

In order to solve the problem (1), a researcher named caoxiang has proposed an on-line adaptive framework, and a controlled experiment has validated its advantages. But the framework has rooms to be improved and expanded. This pa-

pervisor. His current research interests include real time system, human computer interaction.

per proposes a new framework that improves and expands the framework proposed by caoxiang. The framework puts forward an on-line two-way adaptive strategy that adopts the thought of human to human natural interaction while caoxiang's framework only concentrates on the adaptive function of the systems to users, ignoring users' learning capacity of the systems. Besides the two-way adaptive strategy, the framework also consists of an on-line pen gestures query and help system which solves the problem (2). In the paper, we implements a prototype in which users can manipulate photos by gestures and a controlled comparative experiment validates the framework's advantages over caoxiang's framework in usability.

This work is supported by the National Natural Science Foundation of China under grant No. 60503054/No. U0735004 and the National High Technology Development Program of China under grant No. 2007AA01Z158.

The projects mainly concentrate on proposing effective adaptive algorithms and methods which can be adopted in multimodal applications. So far, some papers have been published in the international proceedings and journals on this topic. The paper's main contribution is providing an new adaptive strategy to pen gestures input that improve the systems' usability.