

## 连续交互空间下的混合手势交互模型\*

刘杰<sup>1,2,3</sup>, 黄进<sup>1</sup>, 田丰<sup>1</sup>, 胡伟平<sup>1</sup>, 戴国忠<sup>1</sup>, 王宏安<sup>1</sup>



<sup>1</sup>(人机交互北京市重点实验室(中国科学院 软件研究所),北京 100190)

<sup>2</sup>(中国科学院大学,北京 100190)

<sup>3</sup>(桂林电子科技大学 计算机与信息安全学院,广西 桂林 541004)

通讯作者: 田丰, E-mail: tianfeng@iscas.ac.cn

**摘要:** 分析了触控交互技术在移动手持设备及可穿戴设备的应用现状及存在的问题.基于交互动作的时间连续性 & 空间连续性,提出了将触控交互动作的接触面轨迹与空间轨迹相结合,同时具有空中手势及触控手势的特性及优点的混合手势输入方法.基于连续交互空间的概念,将混合交互手势、空中手势、表面触控手势进行统一,建立了包括空中层、表面层、混合层的连续交互空间分层处理模型.给出了统一的信息数据定义及数转换流程.构建了通用性的手势识别框架,并对轨迹切分方法及手势分类识别方法进行了阐述.最后设计了应用实例,通过实验,对混合交互手势的可用性 & 连续空间分层处理模型的可行性进行了验证.实验结果表明,混合手势输入方式同时兼具了表面触控输入及空中手势输入的优点,在兼顾识别效率的同时,具有较好的空间自由度.

**关键词:** 人机交互;手势交互;交互模型;交互空间;混合手势

中图分类号: TP391

中文引用格式: 刘杰,黄进,田丰,胡伟平,戴国忠,王宏安.连续交互空间下的混合手势交互模型.软件学报,2017,28(8):2080-2095. <http://www.jos.org.cn/1000-9825/5123.htm>

英文引用格式: Liu J, Huang J, Tian F, Hu WP, Dai GZ, Wang HA. Hybrid gesture interaction model in the continuous interaction space. Ruan Jian Xue Bao/Journal of Software, 2017,28(8):2080-2095 (in Chinese). <http://www.jos.org.cn/1000-9825/5123.htm>

## Hybrid Gesture Interaction Model in the Continuous Interaction Space

LIU Jie<sup>1,2,3</sup>, HUANG Jin<sup>1</sup>, TIAN Feng<sup>1</sup>, HU Wei-Ping<sup>1</sup>, DAI Guo-Zhong<sup>1</sup>, WANG Hong-An<sup>1</sup>

<sup>1</sup>(Beijing Key Laboratory of Human-Computer Interaction (Institute of Software, The Chinese Academy of Sciences), Beijing 100190, China)

<sup>2</sup>(University of Chinese Academy of Sciences, Beijing 100190, China)

<sup>3</sup>(School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin 541004, China)

**Abstract:** This article analyzes the application and existing problems of touch interaction for mobile handsets and wearable device. Based on time continuity and space continuity of interactive action, a hybrid gesture input method with characteristic and advantages of air gesture and touch gestures is proposed by combining the contact trajectory and in-air trajectory of the interactive action. Based on the concept of continuous interaction space, hybrid interaction gestures, air gestures, and surface touch gestures are unified to establish a continuous interaction space hierarchical processing model including the air layer, surface layer and hybrid layer. The definition of

\* 基金项目: 国家重点研发计划(2016YFB1001402); 国家自然科学基金(61422212, 61232013, 61170182, 61273269); 国家高技术研究发展计划(863)(2015AA020506, 2015AA016305)

Foundation item: National Key Research and Development Program of China (2016YFB1001402); National Natural Science Foundation of China (61422212, 61232013, 61170182, 61273269); National High-Tech R&D Program of China (863) (2015AA020506, 2015AA016305)

收稿时间: 2015-05-20; 修改时间: 2015-09-10; 采用时间: 2015-09-17; jos 在线出版时间: 2016-10-11

CNKI 网络优先出版: 2016-10-12 16:26:34, <http://www.cnki.net/kcms/detail/11.2560.TP.20161012.1626.002.html>

information data and the transformation process are then provided. The general gesture recognition framework is also constructed. Furthermore, the trajectory segmentation method and gestures classification methods are introduced in detail. Finally, an application instance is designed, and the availability of the hybrid interaction gestures and feasibility of the continuous interaction space hierarchical processing model are verified. Experimental results show that the presented hybrid gesture interaction model possesses advantages of both air gesture and touch gestures, and has good recognition efficiency and better spatial freedom.

**Key words:** human-computer interaction; gesture interaction; interaction model; interaction space; hybrid gesture

触控技术是指不需要实体按键,直接使用手指或笔尖接触设备表面来进行交互的技术<sup>[1,2]</sup>。近年来,随着手机、平板电脑、笔记本电脑、可穿戴设备等各种移动设备的普及,基于手指触控输入方式已经成为许多移动计算设备,如手机、平板、笔记本电脑、可穿戴设备的主要输入方式<sup>[3-7]</sup>,成为与键盘、鼠标并立的主流交互方式<sup>[8]</sup>。目前,大多数的移动设备、可穿戴设备都使用平面触控屏幕,基于二维平面的触控手势<sup>[9]</sup>来进行触控交互。

与传统的按键输入方式相比,触控交互方式更为精确、简便、自然、容易使用,可以提供更为丰富的交互动作,如旋转、缩放等<sup>[10]</sup>,但是也存在一些局限性。

- 可触控面积受到设备体积的限制。对于手持移动设备或可穿戴设备而言,由于对便携性的特殊要求,其体积一般都较小,因此可触控表面积受到外形尺寸的天然限制<sup>[11]</sup>。这一问题对可穿戴设备而言尤为突出。若可触控面积过小,“胖手指”<sup>[12]</sup>的问题就会凸显,对交互效果产生较大的影响<sup>[6]</sup>,甚至无法进行有效的多点触控操作。
- 交互动作被限制在触控表面上。由于触控交互需要通过手指接触的方式进行交互指令的输入,因此将交互空间限制在触控表面上,交互动作从始至终都被限定在触控平面上进行,这在某种程度上降低了交互效率。例如文本选择或圈定截屏范围,整个选择过程都不能脱离触控表面,一旦手指有短暂的脱离,则会导致输入被中断或甚至重置;同时,在一些非平面的触控表面,如曲面屏幕或柔性屏幕,触控表面的形状和物理特性也会影响触控操作的适用性。在进行触控操作时,其特殊的表面形态有可能会导手指间歇性地脱离触控表面,导致交互效率下降,并影响交互体验。
- 未能充分利用手势的空间信息。从交互的整个过程来看,大部分的触控交互动作都包含从空中开始-接触交互表面-返回空中的过程。而触控交互表面仅能感知到接触时交互表面上的动作信息,截取了落在交互表面上的局部动作轨迹作为交互输入,而未能利用近表面空间动作轨迹信息来辅助触控交互。因此,当接触部分的轨迹过短或有间断时,有可能会导输入识别错误。另外,在进行一些简单手势操作时,如滑动操作,由于无法获取和利用空间中的动作信息,需要在触控表面上延长滑动距离,以获取足够的速度及方向信息来达到所需的滑动效果。
- 与触控手势相比,三维的空中手势交互虽然带来了更大的交互空间,但是远距离、大范围的空中手势精确识别通常需要复杂的软硬件系统,目前还难以应用于移动手持设备及可穿戴设备。同时,空中手势的识别的一些关键性问题,如有意动作及无意动作<sup>[13]</sup>的区分,目前仍没有较为完善的解决方法<sup>[4]</sup>。另外,空中手势缺乏触觉反馈,长时间悬空操作容易感觉疲劳,且难以保持精度。

因此,针对触控交互的局限性,需要突破触控表面对交互动作的限制,充分利用触控手势的近表面空间三维空间信息,将交互空间由传统的二维表面扩展到近交互平面的三维空间中。本文基于交互手势的时间连续性与空间连续性,将触控表面手势延伸到三维空间之中,形成一种同时包含表面轨迹与空间轨迹的混合交互手势。在此基础上,将表面手势、空间手势、混合手势的交互空间进行统一,构建了连续交互空间的分层处理模型,给出了统一的交互信息定义及转换流程;设计统一的手势识别框架,对轨迹切分方法及手势分类识别方法进行了阐述。最后设计了应用实例,对混合交互手势的可用性及连续交互空间的分层处理模型的可行性进行了验证。

## 1 相关研究

基于手势空间维度来划分,可以将手势分为二维手势和三维手势<sup>[14,15]</sup>。二维手势主要是基于手部在二维平面上的位置和运动来进行交互操作,而三维手势则将手势动作拓展到了三维空间。目前,基于触控技术的二维手

势已经在可移动设备如手机、平板电脑等得到了广泛的应用<sup>[16]</sup>。与二维手势相比,三维手势识别在硬件和软件上都更为复杂,目前主要应用在更为复杂的交互系统中,并且正逐步向手持移动设备扩展迁移。目前,已经出现了一些可以不接触触控平面的,在近表面空间内(1厘米左右)进行“悬浮触控”的交互设备,如三星的 GALAXY S4 手机、索尼的 Xperia MT27i 手机等。

基于触摸输入位置与触摸屏相对位置,可以将二维交互手势的研究分为屏幕表面、背面及侧面这 3 种<sup>[5]</sup>。其主要的相关研究有:美国马里兰大学的 Lai 等人在 CHI 2014 上发表研究,通过在触摸屏上显示一个能够加速移动的虚拟的拇指来解决单手操作时难以触及屏幕远端的问题<sup>[17]</sup>;韩国科学技术学院在 CHI 2014 发表文章,提出使用两次或多次屏幕点击之间的位置距离和时间间隔来扩展输入词汇,形成新的输入手势<sup>[3]</sup>;韩国蔚山科学技术大学的 Oakley 等人在 CHI 2014 上,通过在类似手表的小型交互设备的侧边添加电容传感器,使设备狭窄侧边也能进行触控操作<sup>[18]</sup>;美国微软研究院的 Baudisch 等人在 2009 SIGCHI 上发表研究<sup>[19]</sup>,通过在设备背部安装与屏幕尺寸相同的触摸板,实现触摸及按压等操作;三菱电机美国研究所的 Wigdor 等人于 UIST 2006 发表研究,通过在数字桌面的底面安装触摸传感器,使得数字桌面的底面也具有了触控交互能力<sup>[20]</sup>;谷歌公司的栗阳等人 2014 年在《美国计算机学会人机交互作用汇刊》上发表文章,提出了一种名为 Touch<sup>[21]</sup>的开发工具,提供了常用、通用手势库,供开发者直接调用。开发者也可以通过录制手势或编辑 XML 手势定义文件的方式,快速建立自定义手势。开发者可用通过 Touch 工具提供的 API,将自定义手势集成到自己的应用程序中。Touch 工具将连续的手势动作分解为多个基本动作片段,通过各个片段的手势状态、当前视图状态构成的动态贝叶斯网络模型,计算事件发生的条件概率,依据概率来驱动事件,并调用相应的处理函数。这些方法实际上是针对平面触控手势的平面轨迹部分,对触控动作进行了动作片段划分。在微观上,通过动作片段之间的时间序列关系、应用视图上的空间关系,构造贝叶斯网络模型并计算事件发生概率,从而决定采取何种动作响应。通过各个片段的微观处理和响应,宏观上对片段序列进行了过滤、拼接操作,最终将片段动作组合为一个完整的交互动作。

对于三维空中手势,其交互空间主要位于设备周围的三维空间。目前已有的相关研究包括:东京大学的 Niikura 等人于 SIGGRAPH 2010 发表研究,在手机底部附加一个高帧率的摄像头,捕捉手指在空中的点击动作,可以在手机上方在空中点击屏幕键盘,并提供震动反馈<sup>[22]</sup>;德国电信实验室的 Ketabdar 等人于 IUI 2010 发表研究“MagiTact”,通过在手上佩戴或持握永磁体,使用手机磁场传感器检测手机周围空间磁场变化,以识别空中手势输入动作<sup>[23]</sup>;德国电信实验室的 Kratz 等人于 MobileHCI 2009 发表研究“HoverFlow”,通过在移动设备的边框上附加多个红外接近传感器,可以粗略追踪屏幕上方的手部运动<sup>[24]</sup>;微软研究院剑桥实验室的 Hilliges 等人于 UIST 2009 发表文章,通过深度相机来检测交互桌面上方空间中的手部姿态<sup>[25]</sup>。

传统的手势交互研究主要集中于二维表面手势或空中手势这两个方面<sup>[26]</sup>,实际上将交互空间划分为平面表面和平面周围空间两个部分,交互手势也主要以平面手势和空间手势为主。虽然 Marquardt 等人基于数字桌面应用研究,2011 年在 INTERACT 2011 提出了连续交互空间的概念,将空中手势和触控手势结合起来,但其主要着眼点在于数字桌面平面触控与空间手势交互之间的无缝切换<sup>[26]</sup>。Chen 等人在 UIST 2014 上发表研究,基于移动设备,将平面点击操作与空中手势相结合,使用移动设备触控平面对空中手势进行了切分<sup>[5]</sup>。韩国科学技术学院的 Gu 等人在 CHI 2013 上设计了基于光敏三极管的键盘尺寸大小的触摸板 LongPad,可以进行力度感应、接近感应、触控操作,并设计了将悬空操作、触控操作相结合的虚拟钢琴应用<sup>[27]</sup>。加拿大维多利亚大学的 Moghaddam 在 CHI 2011 上发表研究,使用红外摄像头捕捉交互透明屏幕上方的手部动作,能够区分触摸操作及近表面的空中手势,并进行实时追踪<sup>[28]</sup>。这些研究虽然都涉及到了触控手势、空中手势,但实际上都是以触控交互平面作为分界,把交互空间分为二维平面及三维空间两部分,将平面手势与空间手势作为两种不同的交互方式进行区分对待,而没有将触控操作作为三维交互动作的一部分或子集。

## 2 混合手势交互及连续交互空间

在理想情况下,如果对手指或手部运动的追踪足够精确,其运动轨迹在空间中应是一条时间、空间上连续的曲线,应具有时间连续性、空间连续性。

- 时间连续性:进入交互空间后,对应任何一个时间点,都有相应的轨迹信息.
- 空间连续性:手势轨迹是一条连续的三维空间曲线,没有间断或跃迁.

从时间连续性及空间连续性来看,交互表面的触控动作实际上是整个三维空间动作的一个部分或子集,在进行动作识别和理解时,应该将 1 段或多段交互表面上的轨迹作为整个三维空间的连续交互动作的组成部分来进行整体识别和理解.对表面触控动作的理解和识别不应仅仅依赖于交互表面上的二维信息,应充分考量平面接触前、接触时、接触后整个过程来进行理解和交互.因此,基于时间、空间连续性,我们把表面触控视为三维交互动作的 1 个或多个组成部分,将交互动作的表面轨迹部分与空中轨迹部分结合到一起,形成一种介于表面手势与空间手势之间的混合交互手势(如图 1 所示).



Fig.1 Mixed gesture sample  
图 1 混合手势示例

基于手势轨迹与设备交互表面的关系,我们将交互空间划分为 3 层:空中交互层、表面交互层、混合交互层(如图 2 所示).

- 空中交互层的物理空间位于设备的交互表面以上,主要通过空中手势的方式进行输入.
- 表面交互层的物理空间为设备的触控交互表面本身,主要通过直接接触的单点或多点触控方式进行交互.
- 混合交互层的物理空间包括交互表面及附近的空间,因此其交互手段以结合了空间及表面手势的混合交互手势为主.

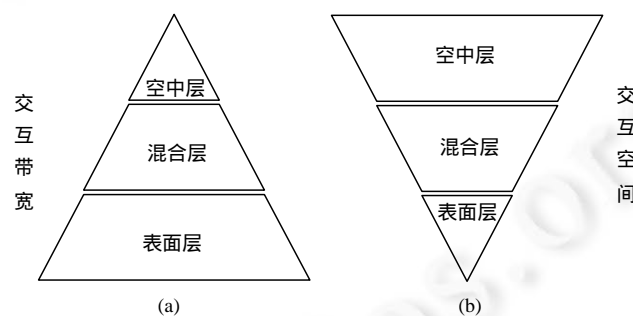


Fig.2 Composition of continuous interactive space  
图 2 连续交互空间的组成

在大多数情况下,随着与传感器距离的增加,其信息带宽会逐渐减少.信息带宽指各个交互通道的有效交互数据流量的带宽之和<sup>[14,29]</sup>,受采样精度、采样速率、传输速率的限制和影响.对于目前的大多数移动设备而言,距离其交互表面距离越远,其交互能力越弱,并有可能带来错误率的上升<sup>[30]</sup>.以具有“悬浮触控”功能的移动设备的电容感应屏幕为例,使用互电容技术可以在屏幕表面实现多点触控,有较好的触控交互精度和流畅度;离开屏幕表面的近表面空间,则需要使用自电容感应技术,只能感应到单点,且精度和实时性大为降低.因此,对于一般的移动、可穿戴设备的传感器来说,以其触控交互表面为基准,距离其交互表面越远,有效的交互信息带宽就越小.由于表面交互层可以包括压力、接触面积、多指识别等多种信息,且具有较高的精度,因此其总带宽最大;混

合交互层包含了交互表面及附近的空间,需要在二者之间进行多次切换,因此总交互带宽次之;空中交互层的有效交互带宽最小,如图 2(a)所示.

由于移动设备的物理尺寸的限制,其交互表面尺寸有限,触控交互方式天然地受到交互表面尺寸的制约;而脱离表面的交互方式可以带来更大的交互空间及自由度.就物理交互空间而言,表面交互层为触控表面本身,混合交互层包含了触控表面及其上方的近表面空间,而空中交互层则为触控表面周围的物理空间,因此其物理交互空间逐层增大(如图 2(b)所示).

由于混合层交互空间包含了交互表面及近表面空间,因此同时具备了二者的优点.在交互带宽方面,混合层优于空中层,同时有接近于表面层的交互带宽及精度(大部分动作轨迹位于交互表面时).由于交互动作可以部分脱离交互表面,因此可以获得比表面层更大的交互空间及交互自由度.

### 3 连续交互空间的分层处理模型

本文提出了一种基于连续交互空间的手势处理层次模型.它在连续交互空间内对空中手势、表面手势、混合手势进行统一的描述和处理.该层次模型如图 3 所示.

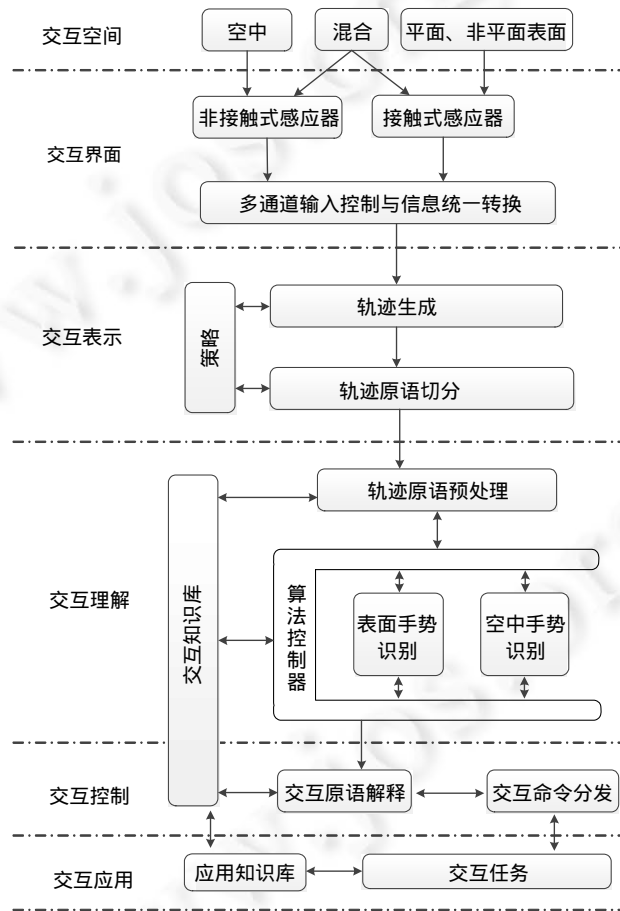


Fig.3 Layered processing model for continuous interactive space

图 3 连续交互空间的分层处理模型

- 层次模型基于统一连续的交互空间,包含了空中交互空间、表面交互空间及混合交互空间,分别对应空中手势输入、表面触控手势输入、混合手势输入.其中,表面交互空间包括平面、曲面及可变表面.

- 交互界面层:负责交互输入感知和交互输出反馈.综合应用多种传感器,感应不同层次的手势输入实时空间位置点信息.对于同一手势输入,有可能会同时被多个层次传感器感应到,因此需要在不同传感器的感知数据流之间进行同步和匹配,包括时间同步和空间匹配.同时,对不同层次空间的手势输入位置点感应信息进行统一描述和转换.
- 交互表示层:将交互界面层感知的实时空间位置点信息进行分组和缓存,形成手势输入的轨迹信息,基于一定的内置策略,将轨迹切分为轨迹原语.
- 交互理解层:对切分后的轨迹进行分类、装配、变换及识别.首先识别手势输入的类型;然后根据类型的不同,将对应的分段轨迹进行装配连接,根据交互应用的属性(二维表面交互任务、三维交互任务)进行轨迹的表面投影或三维变换操作;最后,调用相应的识别算法进行手势识别.
- 交互控制层:将已识别的手势基本动作转换为交互原语,通过应用知识库和交互知识库,将应用层注册的功能集合与交互知识库所提供的基本动作集合进行映射,形成交互原语、复合交互原语,并分发给对应的交互应用.
- 交互应用层:运行各种交互应用程序,最终接收手势输入并进行相应的逻辑处理与反馈.

## 4 连续交互空间的手势输入整合方法

### 4.1 数据描述定义及转换

在连续交互空间分层模型中,将交互表面及其上方的空间视为统一连续的交互空间,其输入包含空中手势、表面手势及混合手势这 3 种方式.在对输入手势信息进行识别和处理时,需要使用统一的动作描述符号,定义如下.

- 位置信息:在交互界面层,由接触式或非接触式的传感器获取到手部或手指的相关位置信息后,传递给多通道输入控制与信息统一转换模块,将各个通道的输入信息统一转换为空间位置信息,记录在当前时间点的输入位置及相关属性.空间位置点有可能位于交互表面上或位于交互表面上方的空间中,表示为一个五元组  $P(Attr, Pos3, Time, Press, Size)$ .其中,  $Attr$  表示输入通道的属性信息,  $Pos3$  表示三维的空间位置信息,  $Time$  表示时间信息,  $Press$  及  $Size$  表示手指接触输入平面时的压力信息及接触面积信息.相邻的空间位置点之间具有空间相连续性及时间连续性.
- 轨迹:由若干空间相邻、时间相续的位置信息点构成的集合,其相邻的空间位置点应同时具有空间连续性及时间连续性,表示为  $Orb=\{P_t|t=0,1,2,3,\dots\}$ .其中,  $P_t$  表示  $t$  时刻的位置信息.轨迹信息可以切分为 1 个或多个轨迹原语.
- 轨迹原语:是指一个完整的、不可再分割的轨迹,记为  $Tra=\{P_t|t=0,1,2,3,\dots\}$ .其中,  $P_t$  表示  $t$  时刻的位置信息,且对于任意的  $P_t, P_{t+1}$ ,同时满足时间连续性、空间连续性、速度单调性.速度单调性包含两个方面:相邻轨迹点之间的瞬时速度值,随时间增加而单方向变化(增加或减少),且变化量应小于一定阈值,即速度的值应平缓单向变化;同时,相邻轨迹点的瞬时速度方向夹角的绝对值应小于一定的阈值,即瞬时速度的方向应平缓变化.若交互表面记为  $S$ ,所有的  $P_t$  都位于交互表面的轨迹原语称为表面轨迹原语,记为  $Ts=\{P_t|P_t \in S, t=0,1,2,3,\dots\}$ ;所有的  $P_t$  都不在交互表面的轨迹原语称为空中轨迹原语,记为  $Ta=\{P_t|P_t \notin S, t=0,1,2,3,\dots\}$ .
- 连续轨迹原语集合:对于空间轨迹原语  $Tra_i, Tra_{i+1}$ ,若  $Tra_i$  的最后一个轨迹点与  $Ta_{i+1}$  的第 1 个轨迹点之间符合时间连续性、空间连续性,则称  $Tra_i, Tra_{i+1}$  可连接.由可连接的轨迹原语所组成的序列,且相邻的轨迹原语都是可连接的,则将其称为连续轨迹原语集合,记为  $Tc=\{Tra_0, Tra_1, \dots, Tra_m\}$ .若其中所有的轨迹原语都为空中轨迹原语,则记为  $Tac$ .若其中轨迹原语都为表面轨迹原语,则记为  $Tsc$ .
- 基本动作:是指对一个连续轨迹原语集合进行解析而得到的有意义的动作  $Act$ ,其集合记为  $A=\{up, down, move, hold\}$ .即所有的轨迹原语都可以解析为向上移动( $up$ )、向下移动( $down$ )、移动( $move$ )、保持( $hold$ )这 4 种基本动作.这里,向上(向下)移动是指位置点  $P_t$  到参考面的投影距离增加(减少).任何交互

动作都是由多个基本动作集组合而成。

- 交互原语:是指手势解析后所得到的具有动作语义的最底层交互命令.它是交互系统所能识别处理的最小交互动作,具有不可分割性,记为  $Cmd$ .在本模型中归纳了 5 种基本的交互输入原语,其集合记为  $C=\{tap,flick,select,release,move\}$ .通过 5 种基本交互原语的组合,可以构成更为复杂的复合交互原语,记为  $CCmd$ .

数据转换流程如图 4 所示,当物理交互空间有手势输入时,在交互界面层,通过多种传感器检测手势的空间位置信息,包括三维空间位置信息  $Pos3$  及触控表面位置信息  $Pos2$ 、压力信息  $Press$  及接触面积  $Size$ .然后,通过时间同步、空间匹配等方式,将这些信息转换为五元组,由于有可能同时有多个输入点(如双手手势),记为  $P_{ni}(n$  为输入点编号, $t$  为时间).在交互表示层分组缓存后,形成各个输入点的轨迹信息  $Orb_n$ .基于时间连续性(采样点时间间隔)、基于空间连续性(位置)以及基于速度单调性(方向及速度值),将轨迹信息  $Orb_n$  切分为轨迹原语.其对应策略函数分别为  $f_{tim}(),f_{dis}(),f_{dir}()$  等.生成的轨迹原语分为表面轨迹原语序列  $Ts_{ni}(i=1,2,...)$  及空中轨迹原语序列  $Ta_{ni}(i=1,2,...)$  两类.通过交互理解层进行动作识别后,生成相应的动作原语序列  $Act_i(i=1,2,...)$ .最后,通过交互控制层转换为交互原语  $Cmd$ 、复合交互原语  $CCmd$ ,分发给应用程序。

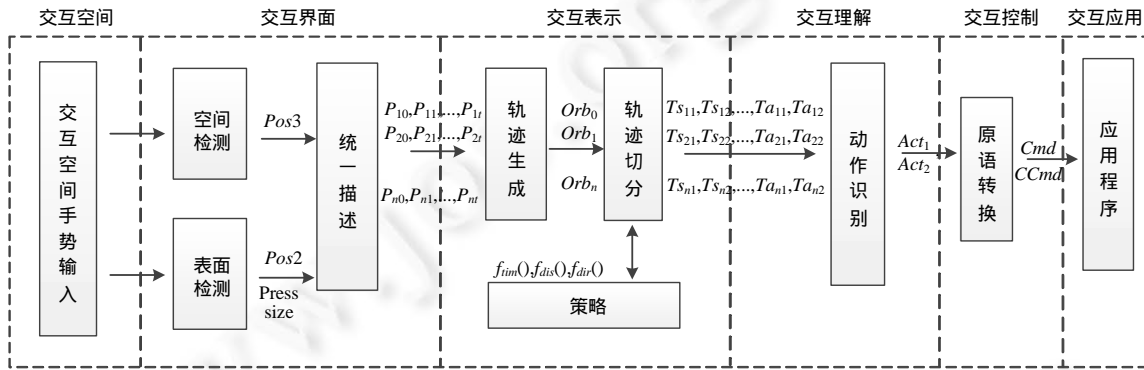


Fig.4 Data conversion between layers

图 4 各层之间的数据转换过程

### 4.2 多参数输入的同步及匹配

由于在统一连续的交互空间内,交互动作包含了空中手势、表面触控手势及混合交互手势,需要使用多种接触式及非接触式传感器进行信息采集,多种传感器数据流之间需要进行时间同步和空间匹配。

- 时间同步

由于系统中可能会同时使用多个接触式、非接触式动作传感器,不同类型的传感器的数据采样频率、数据接口、数据传输协议各不相同,对于同一个手势动作,有可能同时采集到不同类型的动作信息数据,因此需要在不同数据之间进行时间同步.时间同步主要在采样速率、传输速率不同的通道之间进行.可以使用数据插值方法,对低速率的通道进行数据插值,从而使其数据帧率与高速通道的帧率相匹配。

- 空间匹配

在进行动作原始数据采集时,各个传感器一般都会根据自身的参数设定,提供基于自身参考坐标系的动作轨迹原始数据.同时,空间动作传感器提供的是原始三维轨迹信息,而表面触控感应器提供的是原始二维的轨迹信息.因此,需要进行数据维度的统一及空间坐标系的统一.对于空间同步,则主要涉及到同质数据之间的参考坐标系匹配,如空间三维坐标匹配.可以采用设定标定点的方法,获取同一空间标定点在不同传感器坐标系下的位置坐标,求解出坐标系之间的旋转及平移矩阵.例如,可以在屏幕上显示若干个标定点,然后将各个坐标系进行维度统一,将二维坐标变换为三维坐标.通过点击各个标定点的动作,从而获取该标定位置在各个传感器坐标系下的对应空间位置,计算出其他坐标系到屏幕坐标系之间的旋转及平移矩阵,将其他坐标系旋转、平移到屏

幕三维坐标系.

## 5 面向混合手势的通用手势识别框架

### 5.1 通用手势识别框架

在整个层次模型中,最为重要的是负责轨迹切分及动作理解的交互表示层和交互理解层.以这两个层次为核心,其他层次相配合,构成了一个通用的手势识别框架,如图 5 所示.

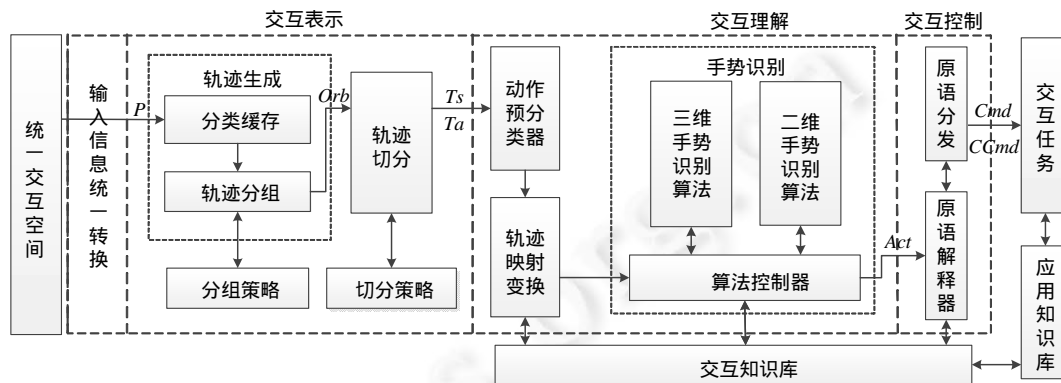


Fig.5 Gesture recognition frame

图 5 手势识别框架

在连续交互空间有动作输入时,其原始的动作空间信息通过交互界面层进行信息统一转换后,形成统一的空间信息描述  $P$ ,传递给交互表示层;然后,根据数据的来源进行分类缓存,形成各个数据源的轨迹信息;然后进行轨迹分组,将不同传感器感应到的、属于同一对象的轨迹信息归为一组,形成对应于每个输入追踪对象的、包含多种数据的轨迹信息  $Orb$ ;随后,基于切分策略对  $Orb$  进行切分,最后形成空中轨迹原语  $Ta$  与表面轨迹原语  $Ts$ .

当轨迹原语  $Ts, Ta$  传递到交互理解层后,首先通过动作预分类器对传入的轨迹原语序列进行初步分类,区分为空中轨迹序列、表面轨迹序列及混合轨迹序列.进入轨迹原语变换模块后,根据应用场景需求进行相应的表面化(表面投影变换)、空间化(空间轨迹构造),也可不做处理,直接传递给手势识别模块.

在完成轨迹原语变换后,输出的空中轨迹、表面轨迹及混合轨迹进入到手势识别模块.针对不同类型的轨迹输入,需要使用相应的识别算法.为了适应多种类型的轨迹及多种形态的移动设备,手势识别模块采用了算法与控制相分离的结构.算法控制器主要负责各算法的控制和调度,根据交互表面的不同形态、不同的输入轨迹类型,调用相对应的手势识别算法进行识别处理.同时,对整个处理过程进行监督和控制,根据识别算法运行过程中的中间状态,对前端的手势预分类器、轨迹原语变换模块进行反馈和调整.手势识别完成后,输入的轨迹原语  $Ta, Ts$  被转换为基本动作  $Act$ .

原语解释器负责对基本动作  $Act$  序列进行解释,转换为交互原语.有两种处理方式:直接转换和复合映射.直接转换是指将单个基本动作  $Act$  直接转换为对应的交互原语  $Cmd$ ;复合映射是将多个  $Act$  构成的序列映射为一个复合交互原语  $CCmd$ .转换完成后,则将交互原语  $Cmd, CCmd$  分发给相应的应用程序.

### 5.2 轨迹切分

由于混合层交互涉及到连续变化(如轨迹)和离散事件(如接触和离开表面),因此使用混合自动机对轨迹切分过程进行描述,如图 6 所示.

混合自动机有 3 个状态  $S_0, S_1, S_2$ , 分别表示无手势轨迹输入、空中轨迹输入、表面轨迹输入.

$S_0$  为初始状态,此时没有检测到输入轨迹点,记为  $P_t = \emptyset$ .  $S_1$  为空中轨迹输入状态,此时,当前轨迹输入点  $P_t$  位



于交互表面上方的空间,即  $P_t \notin S$ .  $P_t$  通过合并操作,并入前一个时刻的空中轨迹原语  $Ta_{t-1}$ ,形成新的原语  $Ta_t$ ,记为  $Ta_t = f(Ta_{t-1}, P_t)$ .  $S_2$  为表面轨迹输入状态,其当前轨迹输入点  $P_t$  位于交互表面内,即  $P_t \in S$ . 通过合并操作  $Ts_t = f(Ts_{t-1}, P_t)$ ,将  $P_t$  并入前一个时刻的表面轨迹原语  $Ts_{t-1}$ ,从而得到当前时刻的轨迹原语  $Ts_t$ .

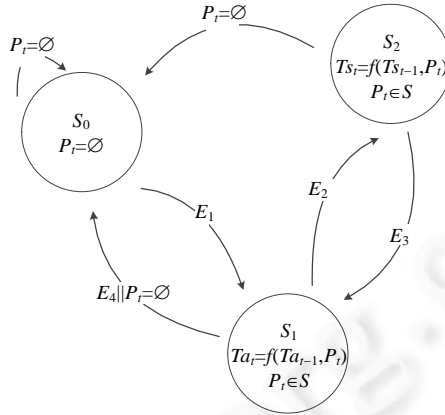


Fig.6 Hybrid automaton for trajectory segmentation  
图 6 轨迹切分的混合自动机

各个状态之间的转换通过一些事件或条件进行控制,如图 7 所示.在初始状态无轨迹输入,因此  $P_t = \emptyset$ ,状态机位于  $S_0$  状态.当开始进行手势交互,手指进入交互空间中时,事件  $E_1$  发生,此时,传感器检测到相关的位置信息  $P_t(Attr, Pos3(x,y,z), Time, Press, Size)$ ,并计算出输入轨迹点到交互表面的投影距离  $H$ .此时,状态机转入  $S_1$  状态,  $P_t \in S$ .当手指向下运动,接触交互表面时,事件  $E_2$  发生,此时,轨迹点属性  $Attr$  发生改变,投影高度  $H$  跃变为 0(有可能无法精确追踪到手指与表面的初始接触点),而交互表面压力值  $Press$ 、接触面积  $Size$  发生跃变.此时,  $P_t \in S$ ,进入  $S_2$  状态.当手指离开交互表面时,事件  $E_3$  发生,压力值  $Press$ 、接触面积  $Size$  跃变为 0,而投影高度  $H$  也发生跃变,此时进入  $S_1$  状态,  $P_t \in S$ .当手指继续抬升,达到交互空间边缘时,事件  $E_4$  发生,由于无法检测到手势输入,此时  $P_t = \emptyset$ ,进入  $S_0$  状态.为了保证状态机的可靠性,规定:在  $S_1, S_2$  状态时,若出现  $P_t = \emptyset$ ,则返回初始状态  $S_0$ .

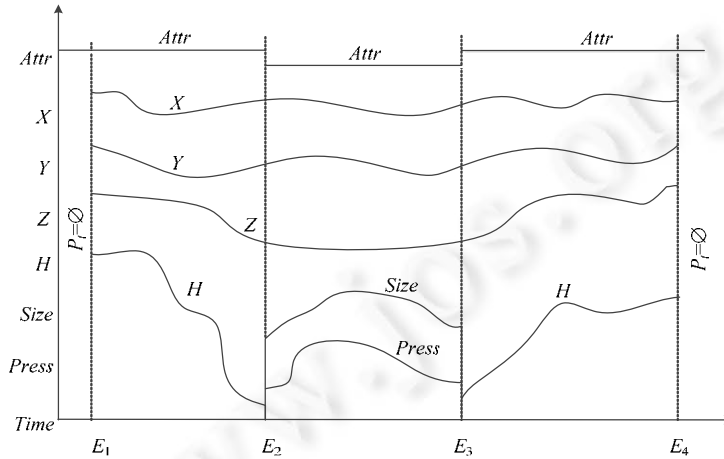


Fig.7 Interactive temporal model  
图 7 交互动作时序模型

5.3 轨迹切分策略

混合自动机处理过程,实质上是将轨迹初步切分为空中轨迹及表面轨迹两类,在  $S_1, S_2$  状态中,需要对轨迹进

行二次切分,以形成轨迹原语  $Ta, Ts$ . 针对平面轨迹,参考文献[31],使用二维平面的轨迹点位置、运动方向、速度值及速度矢量这 4 个属性进行轨迹切分<sup>[31]</sup>. 对于前 3 种属性,采用了限定轨迹分段最大长度(半径小于一定阈值)、限定运动方向角范围(角度小于一定阈值,如  $\pi/2$ )、速度值单调变化(单调递增、单调递减)的切分策略. 对于速度矢量属性,则采用了前 3 个切分策略相结合的方式. 本文采用相似策略,基于时间连续性(采样点时间间隔)、基于空间连续性(位置)以及基于速度单调性(方向及速度值),将切分对象由二维平面轨迹扩展到三维空间轨迹. 其主要的切分策略如下.

(1) 基于时间连续性,将出现时间间隔发生跃变的轨迹点作为切分点. 两个轨迹点的采样时间间隔  $time_{t+1} - time_t$ , 应处于区间  $[1/f_{\max}, 1/f_{\min}]$  内,即采样时间间隔减去区间中点  $(0.5 \times (1/f_{\max} + 1/f_{\min}))$  的差值,应不大于区间长度的  $1/2(0.5 \times (1/f_{\min} - 1/f_{\max}))$ . 由此可以得到评价函数,如公式(1):

$$f_{tim}(P_t, P_{t+1}) = \left| time_{t+1} - time_t - \frac{f_{\max} + f_{\min}}{2f_{\max}f_{\min}} \right| - \frac{f_{\max} - f_{\min}}{2f_{\max}f_{\min}} \quad (1)$$

其中  $f_{\max}$  及  $f_{\min}$  是所允许的传感器最大、最小采样频率.  $time_t$  及  $time_{t+1}$  为相邻的两个空间信息点  $P_t$  及  $P_{t+1}$  的时间戳. 当函数值大于 0 时,相邻轨迹点的时间间隔过大,该点为切分点.

(2) 基于空间连续性,将出现空间坐标跃变的轨迹点作为切分点. 其评价函数如公式(2)所示.

$$f_{dis}(P_t, P_{t+1}) = \sqrt{(x_{t+1} - x_t)^2 + (y_{t+1} - y_t)^2 + (z_{t+1} - z_t)^2} - (time_{t+1} - time_t)v_{\max} \quad (2)$$

其中  $P_t$  及  $P_{t+1}$  是两个相邻轨迹点的位置信息五元组,其三维空间信息为  $(x_t, y_t, z_t)$  及  $(x_{t+1}, y_{t+1}, z_{t+1})$ ;  $time_t$  及  $time_{t+1}$  为时间戳;  $v_{\max}$  是一般情况下手势轨迹所能达到的最大瞬时速度. 当相邻轨迹点之间的位移速度过大,导致位移值大于阈值(取决于  $v_{\max}$ )时,评价函数的值大于 0,则该点作为切分点.

(3) 基于速度单调性,相邻轨迹点的瞬时速度和方向变化较为单调平缓变化,不能有突然的跃变. 因此,将瞬时速度大小或方向有较大变化,或速度变化的单调性发生改变的轨迹点作为切分点.

- 对于轨迹点,其瞬时速度的计算公式为

$$V(P_{t+1}) = \frac{\sqrt{(x_{t+1} - x_t)^2 + (y_{t+1} - y_t)^2 + (z_{t+1} - z_t)^2}}{time_{t+1} - time_t} \quad (3)$$

- 基于速度变化的单调性评价公式为

$$f_{mon}(P_{t+1}) = \frac{V(P_{t+1}) - V(P_t)}{V(P_t) - V(P_{t-1})} \quad (4)$$

当值大于 0 时,该点出现了速度变化的拐点,因此作为切分点.

- 基于速度值变化率的评价公式为

$$f_{val}(P_{t+1}) = \frac{V(P_{t+1}) - V(P_t)}{time_{t+1} - time_t} - a_{\max} \quad (5)$$

当值大于 0 时,速度变化率大于阈值  $a_{\max}$ ,则该轨迹点作为切分点.

- 基于速度方向变化的评价公式为

$$f_{dir}(P_{t+1}) = \left| \frac{x_{t+1}x_t + y_{t+1}y_t + z_{t+1}z_t}{\sqrt{x_{t+1}^2 + y_{t+1}^2 + z_{t+1}^2} \sqrt{x_t^2 + y_t^2 + z_t^2}} \right| - c_{\max} \quad (6)$$

当值大于 0 时,相邻轨迹点的瞬时速度方向夹角余弦值大于阈值  $c_{\max}$ ,则该点作为切分点.

#### 5.4 手势分类

为了描述交互系统可以识别的交互手势,需要对手势集进行分类定义. 其分类类型如图 8 所示. 图 8(a)为空中手势,图 8(b)为表面手势,图 8(c)~图 8(f)为混合手势. 由于交互空间包含了空中层、表面层、混合层,故对手势集合定义如下:  $G = \{g | g \in Ga \cup Gs \cup Gm\}$ . 其中,  $Ga, Gs, Gm$  分别表示空中手势集合、表面手势集合、混合手势集合. 其中,空中手势集合  $Ga = \{g | g \in Tac\}$ ,表面手势集合  $Gs = \{g | g \in Tsc\}$ . 混合手势由于同时包含了空中轨迹和表面轨迹,且两种轨迹有可能多次交织. 因此只定义了其基本手势类型集  $Gm$ , 包含 4 种基本类型的混合手势(图 8(c)~图

8(f)),由这 4 种基本类型的手势可以组合成更为复杂的混合手势:

$$Gm=\{(g_a,g_s),(g_s,g_a),(g_a,g_s,g_a),(g_s,g_a,g_s)|g_a \in Tac, g_s \in Tsc\}.$$

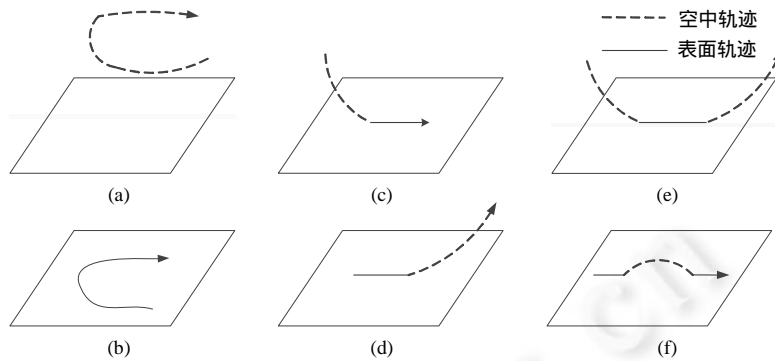


Fig.8 Interactive gesture types  
图 8 交互手势类型

### 5.5 手势预分类器

当轨迹原语  $T_s, T_a$  传递到交互识别层后,首先通过动作预分类器进行初步分类和拼接,其状态转换图如图 9 所示.其中, $S$  为初始状态,其识别缓冲区为空,无轨迹原语.状态  $a \sim$  状态  $f$  分别对应图 8 中的 6 种手势类型.下面以轨迹原语序列  $T_c$  的识别过程为例,若输入的连续轨迹原集合为  $T_c=\{Ta_1, Ta_2, Ts_3, Ts_4, Ta_5\}$ ,在初始状态  $S$  下,首先进入识别缓冲区的为空中轨迹原语  $Ta_1$ ,则进入  $a$  状态,识别为空中手势;后续输入为  $Ta_2$ ,则保持在  $a$  状态,识别为空中手势.输入表面轨迹原语  $Ts_3$  时,则进入  $c$  状态,识别为  $c$  类型的混合手势. $c$  状态下,输入为  $Ts_4$ ,状态保持不变.最后输入为  $Ta_5$ ,进入  $e$  状态,将整个轨迹原语序列  $T_c$  识别为  $e$  类型的混合手势.

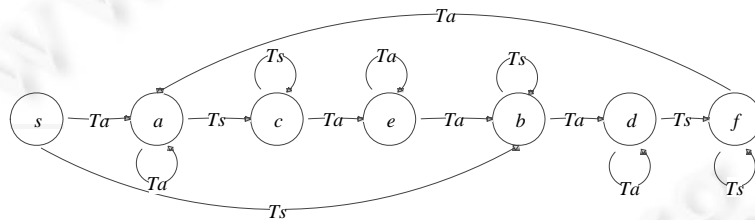


Fig.9 State transition diagram of the type identification for the sequence of the trajectory primitives  
图 9 对轨迹原语序列进行类型识别的状态转换图

### 5.6 轨迹原语变换

通过手势预分类器的处理,输入的轨迹原语序列被识别为 6 种类型,在进行手势识别前,需要根据应用场景与需求对其进行轨迹变换.例如,对于平面的触控屏幕、平面式的 WIMP 范式界面,一般需要将分类后的轨迹原语投影到屏幕平面上,使用平面投影轨迹作为最终输入命令.为了适应多种应用场景需求,将轨迹原语变换处理分为 3 种:直接输出、投影为表面轨迹、构造空间轨迹(如图 10 所示).直接输出是指对轨迹序列不进行变换处理(如图 10(a)所示),直接输出给后继的手势识别模块进行处理;对于只使用交互表面轨迹的场景,则将轨迹序列中的空中轨迹原语  $T_a$  投影到交互表面,得到其表面投影轨迹  $T_s'$ (如图 10(b)所示).经过投影后,轨迹序列只包含表面轨迹原语;对于一些三维交互场景,需要使用交互手势的三维空间信息,则利用序列中的空中轨迹  $T_a$  作为约束条件,将  $T_s$  作为交互表面的投影,构造出其对应的空中轨迹原语  $T_a'$ (如图 10(c)所示).

轨迹原语变换方式的选择主要取决于 3 个因素:当前输入的轨迹原语、当前应用场景状态、之前识别的交互动作.因此可以构造动态贝叶斯网络<sup>[21]</sup>,以轨迹原语变换方式的选择作为响应动作,以前一个交互动作、当前输入的轨迹原语、当前的场景视图状态作为条件,计算响应动作的条件概率,从而依据动态更新的响应动作条

件概率值来选择合适的变换方式.

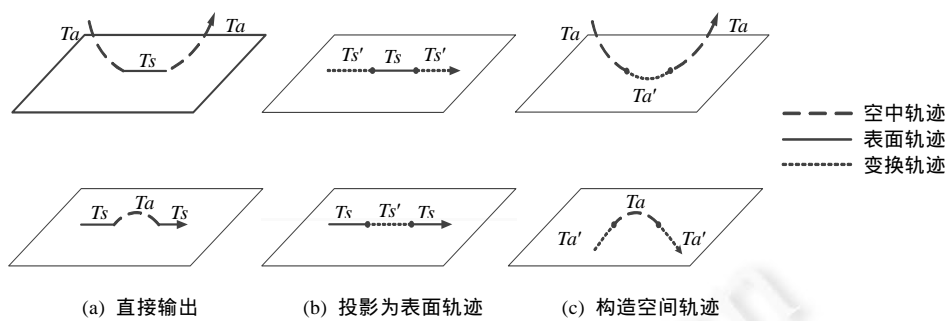


Fig.10 Transformation of the sequence of the trace primitives  
图 10 轨迹原语序列的变换方式

## 6 实验验证

我们设计了一个手势识别的应用实例,来验证空间分层模型的合理性以及手势识别框架模型的可行性.在实验中,我们对连续交互空间层次交互模型进行了简化,主要基于连续的交互空间,使用平面界面应用场景,针对手势识别框架进行实验.

使用三星公司的 GALAXY S4 手机作为主要的实验设备,其触控屏幕使用了自电容感应技术,可以对屏幕上方约 15mm 距离内的单指运动进行平面位置追踪.在识别算法方面,采用 \$1 作为二维平面手势的识别算法. \$1 识别算法是华盛顿大学的 Wobbrock 等人<sup>[32]</sup>在 2007 年提出的二维手势识别方法.该算法通过模板匹配的方式进行单笔划二维手势识别.该算法不需要大样本量的手势模板库,仅需要 1 个手势模板样本即可进行手势匹配识别,简单易用,执行速度快,且具有较高的识别率<sup>[14]</sup>.

在 \$1 的标准手势集中,我们依据从简单到复杂的原则,选择了 caret,check,circle,star 这 4 种具有代表性的手势(如图 11 所示),分别以表面手势、空中手势、混合手势这 3 种方式进行相应的手势输入,并记录完成时间及准确率.

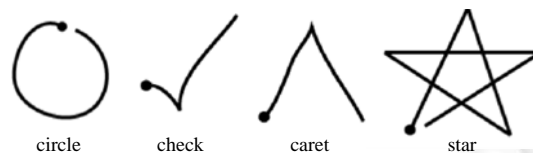


Fig.11 Gestures for experiments  
图 11 实验手势集

本次实验共有 13 名被试者,其中 9 名男性,4 名女性.实验开始前,向被试者介绍实验所使用的 3 种手势输入方式:正常的触控输入方式、非接触的空中输入方式以及允许手指短暂离开交互表面的混合输入方式.用户熟悉操作流程和操作方法后,开始进行实验.对于每种输入方式,进行一组实验,总共 3 组.在每一组实验中,被试者需要使用单指输入 4 种手势,每种手势需要输入 15 次.在实验过程中,要求输入尽可能地快速、流畅.实验完成后,通过问卷调查,让被试者对 3 种输入方式的舒适性、空间自由度程进行打分,分值范围为 1 分~10 分,分值越高,则表明舒适性或自由程度越高.

如图 12、图 13 所示为手势输入所花费的平均时间.由图 12 中可以看出,空中手势输入的时间花费最多,表面触控输入花费时间最少,而混合手势的总时间开销(空中+表面)介于二者之间.混合手势的表面轨迹的时间花费比表面触控手势更小(如图 13 所示).如图 14 所示为输入手势的识别成功率.在 3 种输入方式中,表面触控成功率最高,混合手势次之,空中手势识别率最低.因此,对于 3 种手势输入方式来说,空中输入时间花销最大,准确率

最低;表面触控输入时间花费最小,准确率最高;混合手势输入则介于二者之间。

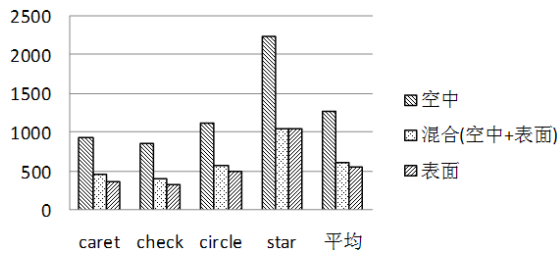


Fig.12 Gesture input time (1)

图 12 手势输入时间(1)

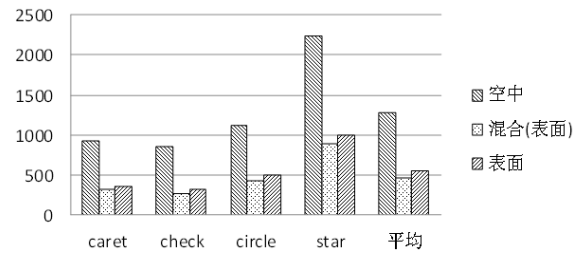


Fig.13 Gesture input time (2)

图 13 手势输入时间(2)

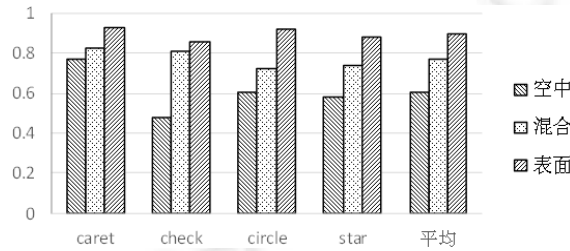


Fig.14 Gesture recognition success rate

图 14 手势识别成功率

对空中手势而言,要达到理想的交互效果,需要大范围的交互空间,而手在空中运动阻力较小(相对于触控而言),运动速度较快.这意味着对空中手势传感器有较高的要求.GALAXY S4 仅能识别屏幕上方向约 15mm 距离内的单指运动二维坐标,且有较为明显的时间延迟(约 0.3s).因此,其识别区间、识别速度都无法达到空中手势的理想要求.在实验中,空中手势输入经常会出现手指移出感应区域而导致识别失败的现象.而明显的延时,使得绘制的轨迹总是滞后于手指的移动,实际上起到了“限速”的作用.同时,基于二维平面坐标,也难以准确识别输入动作的起止点.这些因素的共同作用导致了空中手势输入无论在时间消耗上还是识别成功率上都表现最差.

混合手势输入方式对交互空间的要求相对较低.其输入动作轨迹包含了表面轨迹及空中轨迹(已投影到触控表面),不会由于手指短暂离开交互表面而导致识别失败.同时,可以利用表面轨迹作为约束条件,对其空中轨迹进行平滑、分割及识别.因此,其识别成功率比空中手势高.但是,由于基 GALAXY S4 的自电容传感器仅能识别手指的屏幕投影坐标,而没有空间高度信息,因此,对于起始及退出的空间动作难以进行有效的切分及识别,这对识别成功率有着较为显著的影响.另外,接触表面前及离开表面后的近表面空中轨迹也作为手势输入的有效组成部分,因此与表面触控输入相比,混合手势输入方式的时间开销会更大一些.

实验结束后,针对舒适度及空间自由度对被试者进行了问卷调查.图 15 中显示:表面触控输入的舒适度最高,空中手势输入的舒适度最低.这与手势识别成功率(如图 14 所示)一致,表明手势输入的客观成功率对主观舒适度评价有着重要的影响.图 16 为空间自由度的主观评价,从图中可以看出,混合交互方式空间自由度最高,表面触控方式次之,空中手势输入方式最低.这表明表面触控方式在带来高精度的同时,也损失了部分空间自由度.空中交互方式由于硬件的限制,其交互空间仍较为狭窄(屏幕上方向约 15mm 内的空间);同时,无触觉反馈、难以判断手势输入动作的起止点以及识别滞后等因素,对时间开销及识别率产生了较大的影响,最终也导致了在主观评价中舒适度、空间自由度均为最低.对于混合输入方式,将交互空间由表面扩展到了近表面空间,由于目前硬件实验条件及软件识别算法的限制,识别成功率低于表面触控方式,影响了舒适度评价,但由于交互空间突破了交互表面的限制,得到了三者中最高的空间自由度评分.

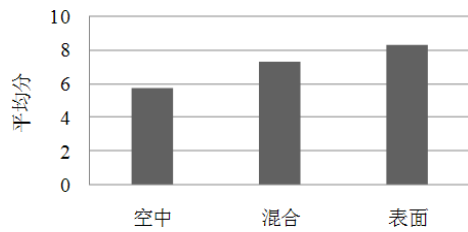


Fig.15 Average comfort degree

图 15 平均舒适度

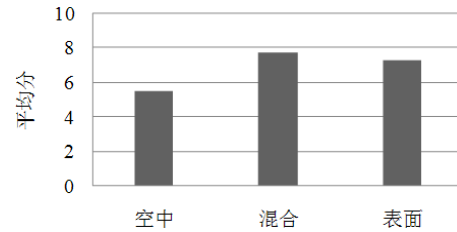


Fig.16 Average space degree of freedom

图 16 平均空间自由度

## 7 总结

本文结合目前交互技术发展趋势,分析了目前触控交互技术在移动机可穿戴设备应用的现状及存在的问题.基于交互动作的时间连续性及空间连续性,提出了将触控交互动作的接触面轨迹与空间轨迹相结合,同时具有空中手势及触控手势的特性及优点的混合手势输入方法.基于连续交互空间的概念,将空中手势、表面触控手势进行统一,建立了包括空中层、表面层、混合层的连续交互空间分层处理模型.详细阐述了连续交互空间的输入整合方法,对各个层次所涉及到的主要数据形式进行了规约和定义,并对数据流的转换过程进行了说明.然后,以分层处理模型中的交互表示层及交互理解层两个层次为核心,构建了通用性的手势识别框架.最后,通过手势识别的应用实验来验证空间分层模型的合理性以及手势识别框架模型的可行性.实验结果表明,混合手势输入方式同时兼具了表面触控输入及空中手势输入的优点,在兼顾识别效率的同时,具有较好的空间自由度.

今后,我们将会围绕连续交互空间,针对混合手势的识别和应用展开进一步研究和实验,研究混合交互手势在其他类型的交互表面,如曲面屏幕、柔性触控表面等的适用性及应用方法.同时,进一步完善手势识别框架,对轨迹切分方法、轨迹切分策略以及算法调度做进一步的研究,以提高对不同类型手势输入、不同类型的交互表面的适应性及识别效率.

致谢 在此,我们向对本文的工作给予支持和建议的中国科学院软件研究所人机交互技术与智能信息处理实验室的苏晓君老师、路璐、陈毅能、贺悦、施大鹏及参加实验的同学表示感谢.

## References:

- [1] Wang F, Cao X, Ren X, Irani P. Detecting and leveraging finger orientation for interaction with direct-touch surfaces. In: Wilson A, ed. Proc. of the 22nd Annual ACM Symp. on User Interface Software and Technology. New York: ACM Press, 2009. 23–32. [doi: 10.1145/1622176.1622182]
- [2] Hourcade JP, Berkel TR. Tap or touch? Pen-based selection accuracy for the young and old. In: Olson G, ed. Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI), 2006 Extended Abstracts on Human Factors in Computing Systems. New York: ACM Press, 2006. 881–886. [doi: 10.1145/1125451.1125623]
- [3] Heo S, Gu J, Lee G. Expanding touch input vocabulary by using consecutive distant taps. In: Jones M, ed. Proc. of the ACM Special Interest Group on Computer-Human Interaction (SIGCHI) Conf. on Human Factors in Computing Systems. New York: ACM Press, 2014. 2597–2606. [doi: 10.1145/2556288.2557234]
- [4] Walter R, Bailly G, Müller J. StrikeAPose: Revealing mid-air gestures on public displays. In: Mackay WE, ed. Proc. of the ACM Special Interest Group on Computer-Human Interaction (SIGCHI) Conf. on Human Factors in Computing Systems. New York: ACM Press, 2013. 841–850. [doi: 10.1145/2470654.2470774]
- [5] Chen XA, Schwarz J, Harrison C, Mankoff J, Hudson SE. Air+Touch: Interweaving touch & in-air gestures. In: Benko R, ed. Proc. of the 27th Annual ACM Symp. on User Interface Software and Technology. New York: ACM Press, 2014. 519–525. [doi: 10.1145/2642918.2647392]
- [6] Bi XJ, Zhai SM. Bayesian touch: A statistical criterion of target selection with finger touch. In: Izadi S, ed. Proc. of the 26th Annual ACM Symp. on User Interface Software and Technology. New York: ACM Press, 2013. 51–60. [doi: 10.1145/2501988.2502058]

- [7] Fukushima S, Kajimoto H. Palm touch panel: providing touch sensation through the device. In: Rekimoto J, ed. Proc. of the ACM Int'l Conf. on Interactive Tabletops and Surfaces. New York: ACM Press, 2011. 79–82. [doi: 10.1145/2076354.2076370]
- [8] Lu L, Tian F, Dai GZ, Wang HA. A study of the multimodal cognition and interaction based on touch, audition and vision. Journal of Computer-Aided Design & Computer Graphics, 2014,26(4):654–661 (in Chinese with English abstract).
- [9] Reisman JL, Davidson PL, Han JY. A screen-space formulation for 2D and 3D direct manipulation. In: Wilson A, ed. Proc. of the 22nd Annual ACM Symp. on User Interface Software and Technology. New York: ACM Press, 2009. 69–78. [doi: 10.1145/1622176.1622190]
- [10] Kane SK, Bigham JP, Wobbrock JO. Slide rule: Making mobile touch screens accessible to blind people using multi-touch interaction techniques. In: Harper S, ed. Proc. of the 10th Int'l ACM SIGACCESS Conf. on Computers and Accessibility. New York: ACM Press, 2008. 73–80. [doi: 10.1145/1414471.1414487]
- [11] Butler A, Izadi S, Hodges S. SideSight: Multi-“touch” interaction around small devices. In: Cousins S, ed. Proc. of the 21st Annual ACM Symp. on User Interface Software and Technology. New York: ACM Press, 2008. 201–204. [doi: 10.1145/1449715.1449746]
- [12] Wigdor D, Williams S, Cronin M, Levy R, White K, Mazeev M, Benko H. Ripples: Utilizing per-contact visualizations to improve user interaction with touch displays. In: Wilson A, ed. Proc. of the 22nd Annual ACM Symp. on User Interface Software and Technology. New York: ACM Press, 2009. 3–12. [doi: 10.1145/1622176.1622180]
- [13] Sohn M, Nam T, Lee W. Designing with unconscious human behaviors for eco-friendly interaction. In: Jr. Olsen DR, ed. Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI), 2009 Extended Abstracts on Human Factors in Computing Systems. New York: ACM Press, 2009. 2651–2654. [doi: 10.1145/1520340.1520375]
- [14] Laviola JJ. An introduction to 3D gestural interfaces. In: Shreiner D, ed. Proc. of the ACM Special Interest Group on Computer Graphics and Interactive Techniques Conf. (SIGGRAPH). New York: ACM Press, 2014. 1–42. [doi: 10.1145/2614028.2615424]
- [15] Konrad T, Demirdjian D, Darrell T. Gesture + play: Full-body interaction for virtual environments. In: Cockton G, ed. Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI), 2003 Extended Abstracts on Human Factors in Computing Systems. New York: ACM Press, 2003. 620–621. [doi: 10.1145/765891.765894]
- [16] Hinckley K, Song H. Sensor synaesthesia: Touch in motion, and motion in touch. In: Tan D, ed. Proc. of the ACM Special Interest Group on Computer-Human Interaction (SIGCHI) Conf. on Human Factors in Computing Systems. New York: ACM Press, 2011. 801–810. [doi: 10.1145/1978942.1979059]
- [17] Lai JW, Zhang DS. ExtendedThumb: A motion-based virtual thumb for improving one-handed target acquisition on touch-screen mobile devices. In: Jones M, ed. Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI), 2014 Extended Abstracts on Human Factors in Computing Systems. New York: ACM Press, 2014. 1825–1830. [doi: 10.1145/2559206.2581158]
- [18] Oakley I, Lee D. Interaction on the edge: Offset sensing for small devices. In: Jones M, ed. Proc. of the ACM Special Interest Group on Computer-Human Interaction (SIGCHI) Conf. on Human Factors in Computing Systems. New York: ACM Press, 2014. 169–178. [doi: 10.1145/2556288.2557138]
- [19] Baudisch P, Chu G. Back-of-Device interaction allows creating very small touch devices. In: Jr. Olsen DR, ed. Proc. of the ACM Special Interest Group on Computer-Human Interaction (SIGCHI) Conf. on Human Factors in Computing Systems. New York: ACM Press, 2009. 1923–1932. [doi: 10.1145/1518701.1518995]
- [20] Wigdor D, Leigh D, Forlines C, Shipman S, Barnwell J, Balakrishnan R, Shen C. Under the table interaction. In: Wellner P, ed. Proc. of the 19th Annual ACM Symp. on User Interface Software and Technology. New York: ACM Press, 2006. 259–268. [doi: 10.1145/1166253.1166294]
- [21] Li Y, Lu H, Zhang H. Optimistic programming of touch interaction. ACM Trans. on Computer-Human Interaction, 2014,21(4): 1–24. [doi: 10.1145/2631914]
- [22] Niikura T, Hirobe Y, Cassinelli A, Watanabe Y, Komuro T, Ishikawa M. In-Air typing interface for mobile devices with vibration feedback. In: Deroose T, ed. Proc. of the ACM Special Interest Group on Computer Graphics and Interactive Techniques Conf. (SIGGRAPH). New York: ACM Press, 2010. [doi: 10.1145/1836821.1836836]
- [23] Ketabdar H, Yüksel K, Roshandel M. MagiTact: Interaction with mobile devices based on compass (magnetic) sensor. In: Rich C, ed. Proc. of the 15th Int'l Conf. on Intelligent User Interfaces. New York: ACM Press, 2010. 413–414. [doi: 10.1145/1719970.1720048]

- [24] Kratz S, Rohs M. HoverFlow: Expanding the design space of around-device interaction. In: Oppermann R, ed. Proc. of the 11th Int'l Conf. on Human-Computer Interaction with Mobile Devices and Services. New York: ACM Press, 2009. 1–8. [doi: 10.1145/1613858.1613864]
- [25] Hilliges O, Izadi S, Wilson AD, Hodges S, Garcia-Mendoza A, Butz A. Interactions in the air: Adding further depth to interactive tabletops. In: Wilson A, ed. Proc. of the 22nd Annual ACM Symp. on User Interface Software and Technology. New York: ACM Press, 2009. 139–148. [doi: 10.1145/1622176.1622203]
- [26] Marquardt N, Jota R, Greenberg S, Jorge JA. The continuous interaction space: Interaction techniques unifying touch and gesture on and above a digital surface. In: Proc. of the 13th IFIP TC 13 Int'l Conf. on Human-Computer Interaction (INTERACT 2011). Berlin: Springer-Verlag Press, 2011. 461–476. [doi: 10.1007/978-3-642-23765-2\_32]
- [27] Gu J, Heo S, Han J, Kim S, Lee G. LongPad: A touchpad using the entire area below the keyboard of a laptop computer. In: Bauersfeld P, ed. Proc. of the ACM Special Interest Group on Computer-Human Interaction (SIGCHI) Conf. on Human Factors in Computing Systems. New York: ACM Press, 2013. 1421–1430. [doi: 10.1145/2470654.2466188]
- [28] Moghaddam AB, Svendsen J, Tory M, Albu AB. Integrating touch and near touch interactions for information visualizations. In: Tan D, ed. Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI), 2011 Extended Abstracts on Human Factors in Computing Systems. New York: ACM Press, 2011. 2347–2352. [doi: 10.1145/1979742.1979934]
- [29] Zhang JC, Fu N, Qiao LY, Peng XY. Investigation of information bandwidth oriented spectrum sensing method. Acta Physica Sinica, 2014,63(3):77–87 (in Chinese with English abstract). [doi: 10.7498/aps.63.030701]
- [30] Pohl H, Murray-Smith R. Focused and casual interactions: Allowing users to vary their level of engagement. In: Mackay WE, ed. Proc. of the ACM Special Interest Group on Computer-Human Interaction (SIGCHI) Conf. on Human Factors in Computing Systems. New York: ACM Press, 2013. 2223–2232. [doi: 10.1145/2470654.2481307]
- [31] Buchin M, Driemel A, van Kreveld M, Sacristán V. An algorithmic framework for segmenting trajectories based on spatio-temporal criteria. In: Agrawal D, ed. Proc. of the 18th SIGSPATIAL Int'l Conf. on Advances in Geographic Information Systems. New York: ACM Press, 2010. 202–211. [doi: 10.1145/1869790.1869821]
- [32] Wobbrock JO, Wilson AD, Li Y. Gestures without libraries, toolkits or training: A \$1 recognizer for user interface prototypes. In: Shen C, ed. Proc. of the 20th Annual ACM Symp. on User Interface Software and Technology. New York: ACM Press, 2007. 159–168. [doi: 10.1145/1294211.1294238]

## 附中文参考文献:

- [8] 路璐,田丰,戴国忠,王宏安.融合触、听、视觉的多通道认知和交互模型.计算机辅助设计与图形学学报,2014,26(4):654–661.
- [29] 张京超,付宁,乔立岩,彭喜元.一种面向信息带宽的频谱感知方法研究.物理学报,2014,63(3):77–87. [doi: 10.7498/aps.63.030701]



刘杰(1981 - ),男,广西崇左人,博士,讲师,CCF 专业会员,主要研究领域为人机交互技术,图形图像处理.



胡伟平(1991 - ),男,硕士,主要研究领域为嵌入式,生理计算.



黄进(1985 - ),男,博士生,CCF 学生会会员,主要研究领域为人机交互技术,图形图像处理.



戴国忠(1944 - ),男,研究员,博士生导师,CCF 高级会员,主要研究领域为人机交互,计算机图形学.



田丰(1976 - ),男,博士,研究员,博士生导师,CCF 高级会员,主要研究领域为人机交互技术,虚拟现实.



王宏安(1963 - ),男,博士,研究员,博士生导师,CCF 高级会员,主要研究领域为实时智能,用户界面.