

基于空中手势的跨屏幕内容分享技术研究*

黄培恺¹, 喻纯¹, 史元春¹

¹(清华大学 计算机科学与技术系, 北京 100084)

通讯作者: 黄培恺, E-mail: hpk711@126.com



摘要: 在多屏幕环境中,实现跨屏幕内容分享的现有技术忽视了对空中手势的利用.为验证空中手势,实现并部署了一套跨屏幕界面共享系统,支持使用空中手势完成屏幕间界面内容的实时分享.根据一项让用户设计手势的调研,实现了“抓-拖拽”和“抓-拉-放”两种空中手势.通过对两种手势进行用户实验发现:用空中手势完成跨屏幕内容分享是新颖、有用、易于掌握的.总体上,“抓-拖拽”手势用户感受更好,但“抓-拉-放”手势对于复杂情景的通用性更好;前者更适合近屏场景,推荐在屏幕相邻且间距不大、操作距离近时使用;后者更适合远屏场景,推荐在目标屏幕间有间隔屏幕或相距较远时使用.因此,将两种手势相结合使用更为合理.

关键词: 空中手势;内容传输;跨屏幕交互;多屏幕;大屏幕
中图法分类号: TP391

中文引用格式: 黄培恺,喻纯,史元春.基于空中手势的跨屏幕内容分享技术研究.软件学报,2019,30(10):2942-2953. <http://www.jos.org.cn/1000-9825/5781.htm>

英文引用格式: Huang PK, Yu C, Shi YC. Research on sharing content between large displays using mid-air gestures. Ruan Jian Xue Bao/Journal of Software, 2019,30(10):2942-2953 (in Chinese). <http://www.jos.org.cn/1000-9825/5781.htm>

Research on Sharing Content between Large Displays Using Mid-air Gestures

HUANG Pei-Kai¹, YU Chun¹, SHI Yuan-Chun¹

¹(Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China)

Abstract: Mid-air gestures have been largely overlooked for transferring contents among multiple screens. For verification, a cross-screen interfaces sharing system is designed, which enables users to move interface content across screens by gestures in real time. According to the user-designed gesture survey, two mid-air gestures, “grab-drag” and “grab-pull-drop”, are implemented. User experiment shows that they are novel, useful, and easy to learn. In general, “grab-drag” provides better user experience, while “grab-pull-drop” is more versatile in complex scenarios. The former gesture is more suitable when the two screens are adjacent and the operation distance is short too; while the latter suits are better in away-from-screen scenarios, when there is a third screen among target screens or screens are widely spaced. Therefore, a combination of the two gestures will be a rational solution.

Key words: mid-air gesture; content transfer; cross-screen interaction; multiple screen; large display

如今在一块公共区域中设置多块电子显示屏的现象愈加常见^[1],具有多显示屏的交互场景也逐渐出现在各种应用场景之中,如智能指挥室、会议室、智能展厅等等.

在一个具有多块显示屏的公共空间中,每块屏幕都有不同的显示任务.不难想象,有时需要把一块屏幕上的显示内容转移到另外一块屏幕上.例如,智能指挥室中各个分屏上显示了不同的参考信息,有时需要将部分信息

* 基金项目: 国家自然科学基金(61572276, 61672314); 国家重点研发计划(2016YFB1001402)

Foundation item: National Nature Science Foundation of China (61572276, 61672314); National Key Research and Development Project of China (2016YFB1001402)

本文由“自然人机交互新进展”专题特约编辑田丰、喻纯推荐.

收稿时间: 2018-08-18; 修改时间: 2018-11-01; 采用时间: 2018-12-25; jos 在线出版时间: 2019-04-29

CNKI 网络优先出版: 2019-04-30 09:19:01, <http://kns.cnki.net/kcms/detail/11.2560.TP.20190430.0918.005.html>

集中显示在主屏幕上。

目前,多屏环境中传输屏幕内容主要依靠与硬件设备配套的切屏软件来完成,每次屏幕内容的切换转移都需要专门操作软件,如果能够做到利用空中手势自由地在屏幕间移动内容,将会带来全新的用户体验。

基于此,我们针对使用空中手势实现屏幕间内容传递进行了一项调研——让用户设计手势,结果主要有“抓-拖拽”和“抓-拉-放”两种手势,我们还专门开发出一套支持空中手势操作的跨设备界面共享系统,从而实现多屏幕间的界面内容实时分享,在此基础上,本文通过用户实验对这两种手势进行了评测,最后给出一个较为合理的解决方案供参考,并对多屏环境中空中手势的设计提出了一些建议。

1 相关工作

1.1 空中手势

空中手势这种交互方式在多屏环境中具有显著优势,手势交流是人的一种本能,频繁应用于日常生活中,易于接受和理解,在多屏环境中显示屏有大有小,一些现有研究表明,与大屏幕交互时采用空中手势,可以在离屏幕一定距离的地方实现交互,从而更好地获取概览信息^[2,3],拥有较好的用户体验^[4],并且还支持用户与大屏幕上任何地方的目标互动^[5],空中手势的交互模式不用佩戴额外设备,不受限制和束缚,可以自由地与屏幕交互^[6],因此,使用空中手势实现屏幕间的内容传输优势显著。

1.2 大屏幕交互

现有研究中暂未有针对使用空中手势进行多屏幕间内容传输的研究,但是针对空中手势与大屏幕交互的研究有很多,这些研究对我们设计跨屏幕内容传输的空中手势提供了一些思路,Yoo 等人的工作^[7]中调查了用户在与大型公共信息显示屏进行交互时,对于空中手势的偏好,结果显示,虽然“指点-停驻”的方法更准确,但是用户倾向于使用“推”的手势进行选择,使用“抓住-拖拽”的方法进行导览,除此之外,也有文章^[6,8,9]讨论过“指点-停驻”手势以及其他空中手势,但是这些研究都是针对单一屏幕或者针对特殊的应用场景,多屏环境与单一大屏幕存在区别,多屏环境中的屏幕分布于空间中的不同位置,屏幕间存在间隔,而单一大屏幕是完整的一块,用户多屏环境中将一块屏幕的内容转移到另一块屏幕的过程是在两个空间上相互独立的显示区域间转移目标内容,目标内容先从一块显示区离开再进入另一块显示区,然而在大屏幕中,移动内容的过程始终是在一片连贯的显示区域中完成的,所以,相比于在单一大屏幕中移动内容,在多屏幕间传输内容的过程更复杂,不能简单地把多块屏幕看成一块大屏幕来处理,因此,对于多屏内容传输的手势还要进一步加以研究,其难点在于找到符合用户心理预期和认知的用于完成跨设备交互的空中手势。

1.3 跨设备内容传输

对于跨设备内容传输的研究有很多,例如在文献^[10]中,作者实现了让电子信息平滑地流转于个人设备、触控桌面和公共大屏幕之间,使设备间的信息共享更加高效,类似的工作还有文献^[11,12]等,但是这些工作并没有引入空中手势研究,也有文章研究通过空中手势实现跨设备内容传输,例如,在 Makela 等人的工作^[4]中,研究了通过空中手势实现大屏幕向手持移动设备传输内容的两种方式:“抓取-拉”和“抓取-放”,其中,“抓取-拉”是指用户选择内容后保持抓握手势向自己拉取一段固定距离;“抓取-放”是指用户通过抓取选定内容,拖拽放置到大屏幕下方对应自己设备的小人图标上实现内容获取,Makela 等人的研究结果表明,“抓取-拉”更适合用户获取内容到自己设备上的场景,文献^[4]关注的场景是大屏幕和移动设备间的内容传输,本质上还是用户与单块屏幕之间的交互过程,但是,多屏幕间内容传输的目标屏幕在远处,传输过程差别较大,因此需考虑的因素也不完全相同。

综上所述,已有研究对使用空中手势进行跨设备屏幕传输的交互方式还没有明确的研究成果,本文旨在研究基于空中手势实现多屏环境中跨屏幕内容传输的解决方案,为越来越普遍的多屏环境中的人机交互设计提供参考。

2 用户设计手势

在设计空中手势时,我们希望了解在多屏环境中,用户是如何理解使用空中手势跨屏幕传输内容,以及用户对于空中手势的心理预期是怎样的.参考文献[13,14]中的工作可以看出,由用户参与设计是很好的办法.

我们据此进行了一项调研——由用户设计手势.调研共有 20 名参与者,其中有 9 名女性.这 20 人的平均年龄为 31.3 岁(sd=11.5).所有参与者均惯用右手,其中 10 名参与者年龄在 18 岁~27 岁,另外 10 名年龄在 27 岁~55 岁.20 名参与者中 9 人了解或使用过 Kinect 等体感设备.20 人中 8 名为在校学生,其余参与者来自各行业,包括老师、互联网行业、金融行业、公务员等,均有接触到多屏环境的机会.

每次访谈首先带参与者接触多屏幕环境,让其了解在屏幕间传输内容的任务.然后让他们自行设计其认为最自然的一种或多种空中手势,并询问他们设计手势动作的灵感来源.

结果显示,参与者提出的空中手势主要集中于两种方式:(1) 抓住内容移动手臂,从一个屏幕直接拖拽内容到另一块屏幕上,本文中简称为“抓-拖拽”手势;(2) 抓住内容向自己提拉一段距离,转向目标屏幕,做放置动作,本文中简称为“抓-拉-放”手势.这 20 人中最终有 14 人提到了第 1 种方法,12 人提到了第 2 种方法,其中有 6 人同时提出这两种方法.这 20 人中有 8 人明确倾向于第 1 种方法,有 6 人明确倾向于第 2 种方法,其余 6 人觉得两种方法均可或者表示不确定,无法给出选择.有 3 名参与者还提出了其他方法.

这一结果和我们原本的预期结果并不一致.在预期结果中,我们认为会有一些参与者提出,选取内容拖放到屏幕上固定的小图标或按钮上的方法,但是实际上只有一个参与者提到了这种方法.通过再次询问,得到的反馈显示,这种方法可以用,但是没有前两种方法直接、自然.另一个预期结果是,应该有一些参与者提出“指点-停驻”的方法,但是实际结果中只有 1 人提到了这一方法,还有一些参与者表示抓取的方式比这种方式体验性更好.

分析调研结果可以推测出,在采用空中手势时,用户看待屏幕中的内容就像看待一个真实的物体一样,所以才提出一些抓握、拖拽、提拉等生活化的手势.据此,我们实现并评测了“抓-拖拽”和“抓-拉-放”两种空中手势.

3 跨设备界面共享系统

3.1 系统简介

跨设备界面共享系统是我们专门为实现多屏环境下的界面内容共享而开发的.系统基于 C# 语言开发,采用 C/S 结构.客户端程序可以通过进程句柄获取到设备所有界面图像,筛选后可以集中显示,如图 1 所示.系统支持各客户端之间实时传输界面内容,从而实现跨屏幕的信息共享.系统对于管理传输内容有一套完整的解决方案,可以在安装客户端的机器间自由分享界面内容.软件除了支持鼠标键盘的操作外,还支持空中手势操作.每一台设备会连接一台 Kinect 设备,用于识别环境中的操作者,接收手势操作信息.每台显示设备会通过配置文件读取其余各屏幕的相对空间位置,用于后续空中手势的识别.

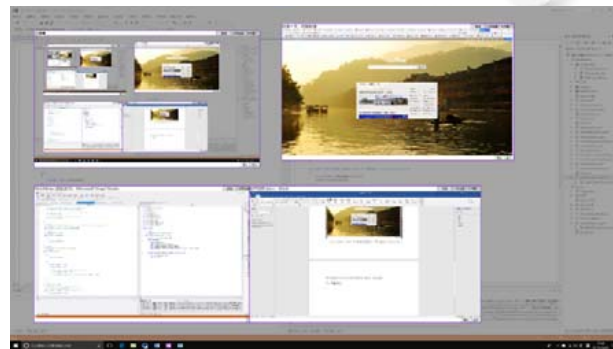


Fig.1 The software operation screenshot

图 1 软件操作截图

3.2 交互锁机制

该系统支持空中手势操作,但一台设备同时只能有一人进行空中手势操作.因为实际应用中一块屏幕前可能会有多名用户,难免会有操作上的冲突,所以为了合理地解决操作资源竞争的问题,我们借鉴了多线程同步的思路,使用锁机制保证每台设备同时只接受一人操作.使用者可以通过特殊动作向设备申请,如果设备处于空闲状态,那么使用者获得操作权,设备上锁,此时的设备只处理该使用者的空中手势.当使用者双手放下时,设备判断使用停止,解锁并释放资源,直到接收到新的申请.如果当前使用者正在操作,则其他人无法取得设备控制权.我们采用的申请动作是双手举过头顶.软件调用 Kincet 提供的接口,获取使用者的骨架空间坐标.令人体部分关节相对于 Kinect 的三维空间坐标为头 T_{head} 、左手 T_{left} 、右手 T_{right} 、身体中心 T_{mid} ,坐标系如图 2 所示,有如下判断.

```
IF  $T_{left}.y > T_{head}.y$  AND  $T_{right}.y > T_{head}.y$ 
THEN 申请控制权
IF  $T_{left}.z > T_{mid}.z + \varphi$  AND  $T_{right}.z > T_{mid}.z + \varphi$ 
THEN 释放控制权
```

当取得控制权后,屏幕的左上角会显示当前控制者的面部图像,光标变为手型,如图 3 所示.

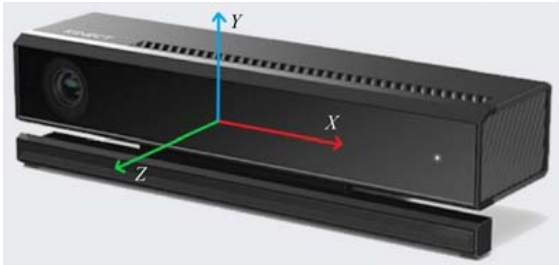


Fig.2 The Kinect coordinate system
图 2 Kinect 坐标系



Fig.3 Interaction with the screen through gestures
图 3 通过手势与屏幕交互

3.3 手位置与光标位置映射

使用者通过空中手势操作设备时,屏幕上的光标显示为手的形状.软件调用 Kincet 提供的接口,获取使用者的骨架空间坐标.首先对头部坐标增加一个偏移量,然后使用偏置后的头部坐标与右手坐标,计算出头手连线,最后计算连线与屏幕所在平面的交点坐标,通过该坐标与屏幕中的位置一一映射,移动右手控制屏幕中光标位置.

令手势指向屏幕对应位置的空间坐标为 T_{point} ,屏幕中心相对于 Kinect 的空间坐标为 T_{screen} ,Kinect 提供的身体中心坐标为 T_{mid} ,头部坐标为 T_{head} .令操作者的中心位置 $T_{center} = T_{head} - (0, 0, \epsilon)$,其中, $\epsilon = T_{head}.z - T_{mid}.z$ (在取得操作权时计算得到,为了避免遮挡干扰,后续使用 T_{center} 而不直接使用 T_{mid}).通过如下计算公式得到:

$$T_{point} = \frac{T_{center}.z - T_{screen}.z}{T_{center}.z - T_{right}.z} (T_{right} - T_{center}) + T_{center}.$$

该公式通过空间坐标运算计算出用户操作中心位置与右手的连线所指向的屏幕位置.令屏幕的分辨率为 $w \times h$,函数 $f(T_{center}, T_{point}, w, h)$ 表示指点位置到屏幕具体像素的映射函数,光标相对于屏幕位置 $(x, y) = f(T_{center}, T_{point}, w, h)$.

操作时使用者只需将手臂前伸,以舒适的姿势在胸前活动,便可以移动光标到屏幕中的任何位置.方法简单,易于掌握,使用效果稳定可靠,在后续的用户实验中反馈较好.本文实现这一方法的主要目的是为了后续实

验能够顺利展开,不影响实验效果,所以不对该方法的交互模型以及疲劳度等相关问题展开讨论.

3.4 界面内容传输

基于前期的调研,我们为实现跨屏幕传输内容设计了“抓-拖拽”和“抓-拉-放”两种手势.这两种手势在使用前都需要先取得大屏幕的控制权.因为所有屏幕的位置信息提前通过配置文件写入系统,所以空中手势的识别均由内容发送设备完成,系统根据手势识别结果发送相应指令到服务器完成内容传输.

“抓-拖拽”手势的整体流程如图 4 所示.首先伸出右手,手掌呈自然张开姿势,同时保证手部可以灵活移动控制光标位置;接下来移动右手位置,将手形光标移动到需要分享的界面内容上,手型由张开状态转变为握拳状态,记录此时的光标坐标为 X_{start} ;最后保持握拳状态,向目标屏幕方向拖拽选中内容,即可实现内容传输.系统根据操作者的拖拽速度和内容拖拽位置来判断是否为拖拽动作,根据拖拽过程中手的位移方向判定目标屏幕.如果需要中途取消操作,只需松开手即可.具体判断规则如下:其中, X_{imp} 表示当前光标位置, T_{point} 表示当前手势指向位置坐标, T'_{point} 表示前一帧数据手势指向位置坐标, $T_{screen1}$ 表示当前屏幕位置, $T_{screen2}$ 表示目标屏幕位置, w 为屏幕的像素宽度, λ 、 θ 、 δ 为常数,可根据实际情况加以调整.

$$\begin{cases} |X_{start} - X_{imp}| \geq \lambda \times w, \\ |T_{point}.x - T'_{point}.x| > \theta, \\ \frac{(T_{point} - T_{screen1}) \cdot (T_{screen2} - T_{screen1})}{|T_{screen2} - T_{screen1}|} \geq \delta |T_{screen2} - T_{screen1}|. \end{cases}$$

判别条件的第 1 条是对屏幕内容被移动的距离进行判断, λ 为常数.判别条件第 2 条的含义是指点位置的移动速度是否达到阈值 θ .第 3 条判别条件判断了从起始屏幕位置到最终手势指向位置的向量在起始屏幕空间位置到目标屏幕空间位置的向量上的投影长度是否足够.

在视觉反馈方面,当抓握住目标内容时,目标内容会被放大,屏幕上显示的其他界面内容均会隐藏,且屏幕背景由半透明灰色变为黄色,抓握处会根据选择的界面内容复制一份半透明内容,示意抓握动作完成.在拖拽过程中,原内容不会移动,复制的半透明内容会随拖拽过程中光标的移动而移动.当判定为拖拽动作后,半透明显示内容会移出屏幕,目标屏幕靠近初始屏幕的一侧会有内容进入动画.



Fig.4 The “grab-drag” gesture

图 4 “抓-拖拽”手势

“抓-拉-放”手势的整体流程如图 5 所示.与“抓-拖拽”手势一样,先伸出右手,保持舒适姿势,手掌自然张开.移动光标到需要分享的界面内容上,手型由张开状态转变为握拳状态,记录此时右手坐标 T_{right1} ,保持握拳状态先

向身体方向拉一段距离,判断规则如下.

$$\begin{cases} |T_{right} - T_{right1}| \geq \alpha(T_{head} \cdot z - T_{mid} \cdot z), \\ T_{right} \cdot z \geq T_{right1} \cdot z + \beta. \end{cases}$$

Kinect 提供的实时身体中心坐标为 T_{mid} , 头部坐标为 T_{head} , 右手坐标为 T_{right} . 公式中第 1 条判别条件的含义是右手抓握后的移动距离超过一定阈值, 该阈值大小根据使用者的身材而发生变化, 身高体长的使用者对应的拉伸距离要更大一些, α 为常量系数. 第 2 条判别条件的含义是使用者握拳拉拽的方向必须是远离屏幕向自己拉近的, β 是常量系数.

记录“拉”动作结束时右手坐标 T_{right2} , 保持握拳向目标屏幕做放置动作, 整个“抓-拉-放”手势完成, 判定规则如下.

$$\begin{cases} |T_{right} - T_{right2}| \geq \alpha(T_{head} \cdot z - T_{mid} \cdot z), \\ T_{right} \cdot z \leq T_{right2} \cdot z - \gamma, \\ \frac{(T_{point} - T_{screen1}) \cdot (T_{screen2} - T_{screen1})}{|T_{screen2} - T_{screen1}|} \geq \delta |T_{screen2} - T_{screen1}|. \end{cases}$$

判别条件第 1 条含义为做“放”动作时, 右手伸出的距离超过一定阈值, 阈值随用户身材变化, 身高体长的用户需要的拉伸距离更大一些. 第 2 条判别条件的含义是使用者“放”动作的手部运动方向是远离身体的. 第 3 条判别条件判断了从起始屏幕位置到最终手势指向位置的向量在起始屏幕空间位置到目标屏幕空间位置的向量上的投影长度是否足够.

设备对判定“拉”动作和“放”动作的参数为 α 、 β 、 γ 、 δ , 可以根据不同情况调整. 同样地, 如果要取消操作, 只需在完成前松手即可.

视觉反馈方面, 当抓握完成后选中的界面内容会放大, 其余内容会被隐藏. 随着拉动作的完成, 选取内容会再次放大片刻以示拉取成功, 随后恢复正常状态. 放置动作完成后, 在目标屏幕上会有内容从起始屏幕方向进入的动画效果.



Fig.5 The “grab-pull-drop” gesture

图 5 “抓-拉-放”手势

除了实现上述两种手势外, 我们还设计了一整套完整的手势交互方案, 保证使用者可以通过手势对系统进行操作, 包括内容位置移动、显示内容关闭、屏幕界面开关、界面内容放大缩小等手势.

4 用户实验

4.1 参与者

为评测两种手势, 我们进行了一项由 20 人参与的用户实验, 实验采用受测者内设计 (within-subject design)

的实验方式.20 名参与者的平均年龄为 29.7 岁(sd=8.15),其中,9 名参与者为在校学生,年龄在 21 岁~29 岁之间,并拥有不同学科背景,其余参与者年龄为 27 岁~50 岁不等,从事各行各业,例如 IT、金融、服务业、自由从业者等.参与者为 11 男 9 女,11 人完全没有 Kinect 或者其他类似空中手势系统的使用经验,其余 9 人中有 8 人只使用过 1~2 次 Kinect,只有 1 人有多于 5 次的使用经验.

4.2 实验环境

实验使用两台屏幕大小为 85 吋(长为 1 868.3cm,高为 1 054.3cm)的一体机.两台机器中间间隔为 50cm,离地高度约 1.2m,各屏幕顶端中部放置一台 Kinect 设备,如图 6 所示.



Fig.6 The experimental screens

图 6 实验屏幕

4.3 实验设计

实验使用跨设备界面共享系统的功能,要求参与者根据指令提示,选取两台设备上的目标界面,通过两种不同的空中手势实现跨设备的界面分享.

在实验开始前,参与者需填写一份统计个人背景信息的简单问卷.在填写完问卷后,向参与者详细介绍项目背景以及完成跨设备界面共享的两种空中手势——“抓-拖拽”手势和“抓-拉-放”手势.参与者随后会观看两种手势的演示动作.在参与者充分了解以上内容后,会根据自己的心理预期分别填写对这两种手势的评价量表(量表内容会在后文中详细介绍).之后,参与者会在专人指导下学习如何使用系统,充分适应用手势控制光标在大屏幕上移动,学会使用抓握动作.最后,通过一个辅助测试确保用户能够正确使用抓握动作在屏幕上准确拖拽界面窗口.

正式实验要求每名参与者按照如下步骤分别实验“抓-拖拽”和“抓-拉-放”两种空中手势:(1) 观看空中手势的使用演示,学习手势动作的基本要领;(2) 学会使用空中手势完成跨设备的界面分享;(3) 按照指令提示选取目标界面内容,使用空中手势完成跨屏幕内容共享;(4) 重复完成以上任务 5 次,每次完成后会稍适休息;(5) 参与者自行体验空中手势,询问感兴趣问题;(6) 填写一份对空中手势体验评价量表.

按照上述过程完成对这两种手势的体验后,参与者需要填写一份比对两种空中手势的问卷.最后,我们会对参与者进行一个简短的访谈.整个实验过程中,我们还在参与者不知情的情况下记录了每名参与者对不同空中手势的学习时间.在自行体验空中手势这一环节,我们根据参与者自行体验的次数、积极性和态度表现等因素,综合评价了参与者对相应空中手势的感兴趣程度.

目前,对于交互设计的主观评估并没有统一的标准,本文研究借鉴了 SUXES 方法^[15],分别收集实验参与者体验前期望与体验后的主观评分,以此评估两种空中手势的体验.前期的期望分数反映了参与者对于新交互方式的先验认知,再通过与体验后的分数比对,可以更好地理解参与者对交互技术的使用体验.

5 实验结果

评分结果采用箱形图绘制,展示了“抓-拖拽”手势和“抓-拉-放”手势的期望评分与体验评分的分布情况.其中,期望评分与体验评分的公共指标有 8 项,如图 7 所示,体验评分有额外的 3 项,如图 8 所示,这些评分取值为

1~7 的整数.数据处理采用箱形图的常规方法,将同一指标的数据分数按照由小到大排列,计算数据的下边缘、下4分位数 Q_1 、中位数、上4分位数 Q_3 、上边缘.其中,下边缘取值为 $Q_1-1.5 \times IQR$,上边缘取值为 $Q_3+1.5 \times IQR$, IQR 为4分位距($IQR=Q_3-Q_1$).

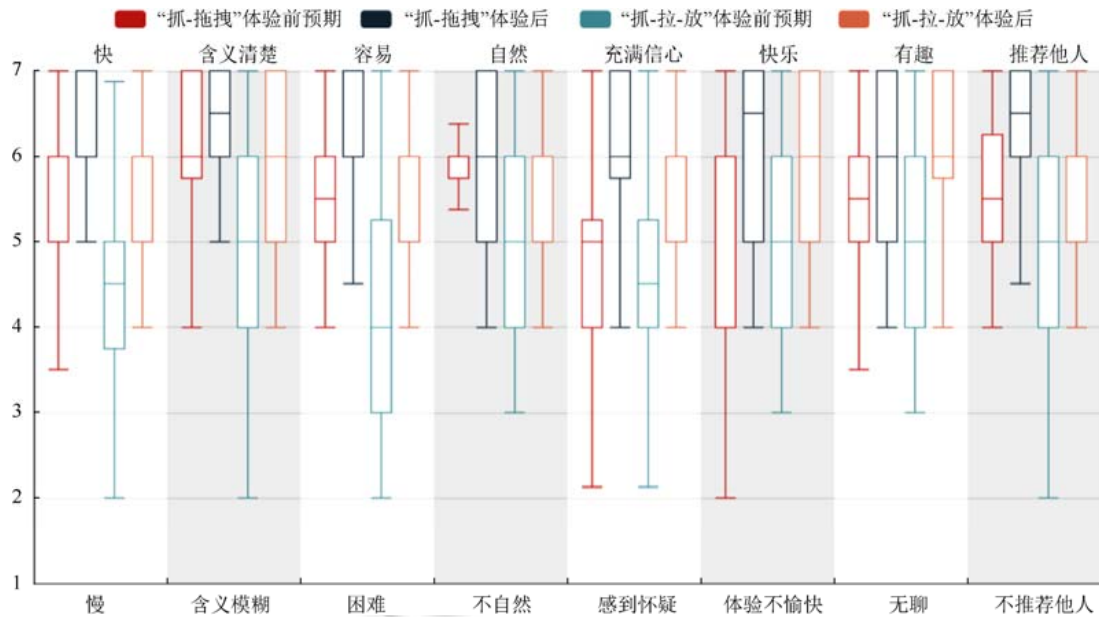


Fig.7 Two gestures before and after scoring box plot

图7 两种手势体验前后评分箱形图

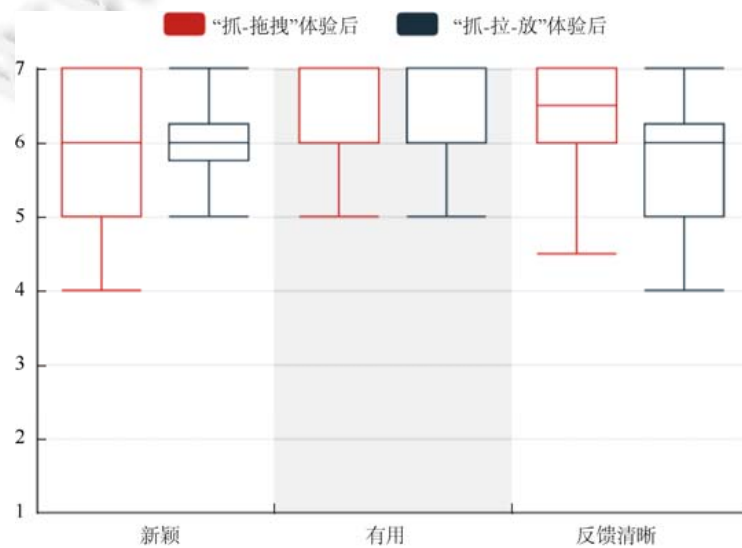


Fig.8 Post-score box plot for two gestures

图8 两种手势体验后评分箱形图

我们对于两种手势的学习速度统计结果见表 1.整体上,两种手势都不难掌握,所有人都能在 3 次以内学会.从学习的次数上看,掌握“抓-拉-放”手势更困难一些.

Table 1 Attempts statistics of two gestures until mastering**表 1** 掌握两种手势尝试次数统计

| 学习次数(次) | 1 | 2 | 3 |
|--------------|---|----|---|
| “抓-拖拽”手势(人) | 8 | 10 | 2 |
| “抓-拉-放”手势(人) | 5 | 10 | 5 |

表 2 记录了参与者在体验时完成 5 次任务过程中的失败次数。可以看出,两种手势的使用出错率都不高。通过检查系统记录下的失败动作日志,发现“抓-拖拽”手势两次失败,其中一次是由 Kinect 识别结果意外抖动导致的,属于特殊情况,还有一次是因为参与者手型识别不准确所致。“抓-拉-放”的失败原因有两个:一是因为手型识别不准确;二是参数设定不合适。出错较多的两名参与者均为女性,动作幅度小于其他参与者。在找到原因后,我们重新调整到合适的参数,再次邀请其中二人尝试,出错率明显降低。

Table 2 Statistics of failures during 5 successful completions**表 2** 成功完成 5 次失败次数统计

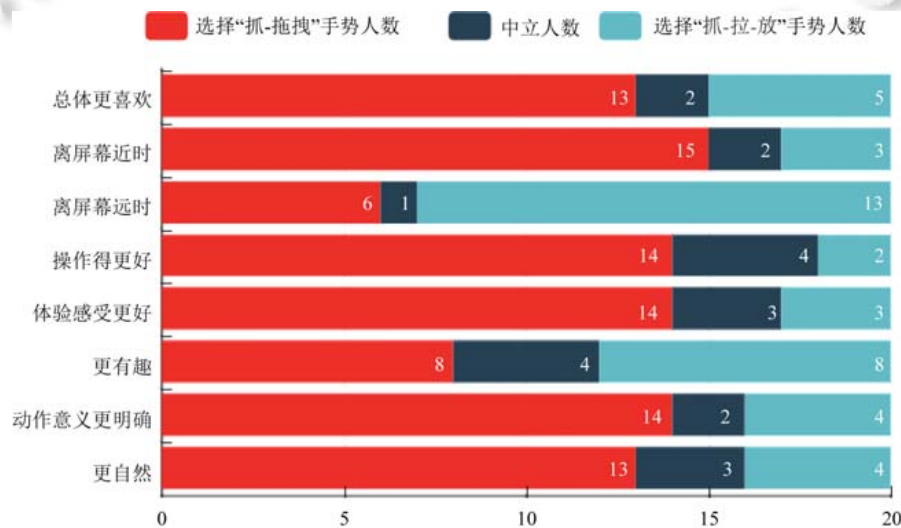
| 失败次数(次) | 0 | 1 | 2 | 3 |
|--------------|----|---|---|---|
| “抓-拖拽”手势(人) | 18 | 2 | 无 | 无 |
| “抓-拉-放”手势(人) | 16 | 2 | 1 | 1 |

完成一次“抓-拖拽”的平均时间为 3 350ms(id=699);完成一次“抓-拉-放”的平均时间为 4 941ms(id=1142)。表 3 记录了每名参与者使用两种空中手势的平均时间消耗,单位为 ms。对两组数据进行方差分析,结果显示出显著差异($F_{1,19}=224.399, P<0.01$)。“抓-拖拽”明显快于“抓-拉-放”。

Table 3 Average time statistics for participants to complete a gesture**表 3** 参与者单次完成手势的平均时间

| 参与者序号 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|-------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| “抓-拖拽”(ms) | 2 781 | 2 835 | 3 539 | 3 987 | 3 429 | 3 422 | 3 401 | 3 525 | 3 798 | 4 256 | 3 293 | 3 735 |
| “抓-拉-放”(ms) | 4 237 | 4 309 | 4 873 | 5 774 | 5 359 | 4 931 | 5 123 | 4 776 | 4 616 | 5 068 | 4 987 | 4 499 |
| 参与者序号 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | | | | |
| “抓-拖拽”(ms) | 3 118 | 2 749 | 3 457 | 3 130 | 3 322 | 3 547 | 3 419 | 3 156 | | | | |
| “抓-拉-放”(ms) | 4 557 | 4 358 | 5 214 | 5 057 | 4 554 | 6 020 | 5 011 | 5 612 | | | | |

对比选择的结果如图 9 所示。

**Fig.9** Participants propensity chart**图 9** 参与者倾向性统计图

6 讨论

整体上看,两种空中手势的评分是正向的.无论“抓-拖拽”手势还是“抓-拉-放”手势,体验后的评分均好于体验前的期望评分.“抓-拖拽”手势的期望评分和体验后评分都要好于“抓-拉-放”手势.

对比体验前与体验后的数据可以发现,在手势速度快慢这一指标上体验后的评分要明显好于体验前的预期.这说明,在速度方面,参与者对空中手势的实现很满意,超出了预期判断.在访谈中参与者谈到为什么体验后认为实际速度比预期要快,主要有两点理由:一是通过抓握手势选取内容的准确度和灵活度比预期要好,可以灵活、快速地移动光标到目标位置,完成抓取动作;二是“拖拽”或者“拉放”的动作比预期要轻松,体验前普遍认为空中手势需要小心翼翼地完成,动作会僵硬拘束.

实验程序也记录下了每种动作的时间开销.由于个体间差异较大,例如对手势使用的熟练度不同、使用手势风格不同(快速尝试或者稳保成功)、个人站位不同、传输内容相对屏幕位置不同等,结果存在差异.但是比对每个人的数据可以发现,结果和评分高低一致,“抓-拖拽”手势快于“抓-拉-放”手势.在访谈中参与者表示,“抓-拖拽”手势在动作分解上可以理解为“抓握”与“拖拽”两个步骤,“抓-拉-放”手势的动作分解是3步;放的动作和拖拽的动作感觉上开销相似,但是相比于“抓-拖拽”,“抓-拉-放”多出了拉的动作,所以在预期评分中会更慢一些.实际体验的评分以及数据统计结果也验证了这一点.

除了速度快慢外,还有第5项指标体验后的评分也明显好于体验前的预期.第5项指标意图了解参与者对自己正确完成空中手势是感到怀疑的还是充满信心的.从预期评分来看,用户对自己能力的信心一般,但是实际体验后信心明显有了提升.这说明,两种空中手势在动作的区分度上得到了参与者的认可,参与者无需担心这两种空中手势会和其他操作动作混淆,例如界面内容位置移动、放大缩小等.对比两种空中手势,参与者对“抓-拖拽”手势的信心更足一些.

在第3项指标(手势动作难易程度)以及第4项指标(手势是否自然)上,两种手势体验后的评分都高于预期,提升幅度较大.这说明,实际体验中的手势动作比参与者们预期的更容易,更自然.原因与前文中提到的类似,在参与者的设想中使用空中手势需要更小心翼翼.

在第8项指标(是否愿意推荐给他人)上,两种手势体验评分都有所提高,“抓-拖拽”手势评分提高得更明显.说明愿意分享“抓-拖拽”手势的人数更多,图9展示的结果也证实了这一点.

第2项指标(手势动作含义是否清楚)和第7项指标(体验是否愉快)的评分体验前后变化差异不大.这也反映出我们对于两种手势的实现符合参与者的想法,也可以理解为我们向参与者描述任务需求背景以及介绍空中手势是客观、准确的.参与者认可实验中的实际体验感受.

在第7项指标(感觉无聊还是有趣)中,“抓-拉-放”手势的分数增长更多.通过对每种手势自由体验环节记录下的体验次数进行分析,也能得出这一结论.参与者体验“抓-拉-放”手势次数更多.观察发现,有些参与者明显对“抓-拉-放”手势思考更多.访谈时我们询问了这些参与者为何看上去对“抓-拉-放”手势更感兴趣,根据回答总结出两方面原因:一是有的参与者很好奇“抓-拉-放”手势在技术上是如何实现的,怎样把方法迁移到不同的场景中;二是一些参与者表示这种手势实现效果高于预期,很有启发性,所以会陷入思考.

再观察两种手势体验后单独打分的3项指标,可以发现,参与者们认为两种方法都很有用,并且很新颖.视觉反馈方面,“抓-拉-放”手势还可以进一步提高.总体上,参与者们都很认可这两种空中手势,并纷纷表示可以用于实际场景,效果令人满意.

观察对比选择的结果,总体上,支持“抓-拖拽”手势的人数明显多于支持“抓-拉-放”手势的人数.在近屏场景中,支持“抓-拖拽”手势的人数更多;但在远屏场景中,支持“抓-拉-放”手势的人更多.访谈中我们也询问了参与者相关原因.因为在近屏场景中,使用“抓-拉-放”手势会不自觉地增大“放”这一动作的幅度(虽然系统并没有做出这样的要求),这就使得动作负担增大;同时,使用“抓-拉-放”手势时注意力需要从移动过程的起始屏幕转移到目标屏幕(并非必然要求),不太适合近屏场景.离屏幕相对较远时,可以更容易获得各屏幕的整体信息,而完成拖拽动作的注意力更多地放在动作起始,所以,注意力平均分配的“抓-拉-放”手势与远屏场景更契合.

在其他指标的对比上(除去有趣一项),支持“抓-拖拽”的人数更多一些,每名参与者也都分别给出了理由.认

为“抓-拖拽”手势体验感更好,操作感更好的参与者普遍认为“抓-拖拽”手势更加简单,上手快;有的参与者认为“抓-拉-放”手势动作意义更明确,而且适用范围更广;还有一些参与者保持中立态度,认为各有优点.有很多名参与者表示,如果可以把两种手势结合起来,根据喜好和不同情景自由选择更好.有6名参与者提到,虽然“抓-拖拽”的体验感很好,但在起始屏幕和目标屏幕中间再增加一块屏幕会怎么样?这种情景下视觉反馈和操作体验都会有变化,通用性显然不如“抓-拉-放”手势.

最后,我们根据参与者的意见,实现了一种结合两种操作手势的方法.在“抓-拉-放”方法保持不变的基础上,新增“捏合-拖拽”的方法代替“抓-拖拽”方法.“捏合-拖拽”使用捏合手势(食指和拇指从张开到捏合)实现界面内容选取,取代了“抓-拖拽”方法中抓握手势的功能.捏合手势会有视觉反馈,捏合处会生成被选界面的半透明复制内容,后续过程与“抓-拖拽”一样.这样可以保证抓握手势只用于界面内容位置调整和“抓-拉-放”方式.“捏合-拖拽”则既保留了“抓-拖拽”的优点,又可以与“抓-拉-放”区分开.在跨屏幕传输内容时,使用者可以根据具体情况选择“捏合-拖拽”或者“抓-拉-放”,不用担心操作产生冲突.最后我们也请参与者们来体验这种结合后的方式,参与者均表示满意.

未来工作可以更进一步地讨论多于两块屏幕情景下的相关问题,甚至还可以探讨更多样环境下的类似问题,例如增加触控桌面等.

7 结 论

我们的工作为多屏环境中的跨屏幕内容传输提供了一种很好的交互方式.实验结果表明,空中手势完成跨屏幕内容传输是新颖、有用、易于掌握的,并且还拥有很好的用户体验.对比两种空中手势“抓-拖拽”和“抓-拉-放”,第1种更适合离屏幕较近的情况,推荐在屏幕相邻且间距不大、操作距离近时使用;第2种正好相反,适合离屏幕较远的情况,推荐在目标屏幕间有间隔屏幕或相距较远时使用.在整体感受上,“抓-拖拽”手势优于“抓-拉-放”手势,不过,“抓-拉-放”手势对于复杂情景的通用性更好.采用两种手势相结合的方式是一种很好的解决办法.

References:

- [1] Nacenta M, Gutwin C, Aliakseyeu D, Subramanian S. There and back again: Cross-display object movement in multi-display environments. *Human-computer Interaction*, 2009,24(1-2):170-229.
- [2] Jakobsen MR, Jansen Y, Boring S, Hornbæk K. Should I stay or should I go? Selecting between touch and mid-air gestures for large-display interaction. In: *Proc. of the Human-computer Interaction—INTERACT 2015, the 15th IFIP TC 13 Int'l Conf.* Springer Int'l Publishing, 2015. 455-473.
- [3] Nancel M, Wagner J, Pietriga E, Chapuis O, Mackay W. Mid-air pan-and-zoom on wall-sized displays. In: *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (CHI 2011)*. ACM, 2011. 177-186.
- [4] Ville M, Jobin J, Tuuli K, Jaakko H, Markku T. It's natural to grab and pull: Retrieving content from large displays using mid-air gestures. *IEEE Pervasive Computing*, 2017,16(3):70-77.
- [5] Vogel D, Balakrishnan R. Distant freehand pointing and clicking on very large, high resolution displays. In: *Proc. of the 18th Annual ACM Symp. on User Interface Software and Technology (UIST 2005)*. ACM, 2005. 33-42.
- [6] Luke H, Martin T, Kazjon G, Anthony C, Judy K. Investigating intuitiveness and effectiveness of gestures for free spatial interaction with large displays. In: *Proc. of the 2012 Int'l Symp. on Pervasive Displays (PerDis 2012)*. ACM, 2012. Article No. 6.
- [7] Yoo S, Parker C, Kay J, Tomitsch M. To DWELL or not to DWELL: An evaluation of mid-air gestures for large information displays. In: *Proc. of the Australian Conf. on Computer-human Interaction (OzCHI 2015)*. ACM, 2015. 187-191.
- [8] Manghisi VM, Uva AE, Fiorentino M, Gattullo M, Boccaccio A, Monno G. Enhancing user engagement through the user centric design of a mid-air gesture-based interface for the navigation of virtual-tours in cultural heritage expositions. *Journal of Cultural Heritage*, 2018,32:186-197.

- [9] Drossis G, Grammenos D, Birliraki C, Stephanidis C. MAGIC: Developing a multimedia gallery supporting mid-air gesture-based interaction and control. In: Proc. of the Int'l Conf. on Human-computer Interaction HCI Int'l 2013. Springer-Verlag, 2013. 303-307.
- [10] Rekimoto J, Saitoh M. Augmented surfaces: A spatially continuous work space for hybrid computing environments. In: Proc. of the SIGCHI Conf. on Human Factors in Computing Systems. ACM, 1999. 378-385. [doi: 10.1145/302979.303113]
- [11] Baudisch P, Cutrell E, Robbins D, Czerwinski M, Tandler P, Bederson B, Zierlinger A. Drag-and-pop and drag-and-pick: Techniques for accessing remote screen content on touch- and pen-operated systems. In: Proc. of the INTERACT 2003. 2003. 57-64.
- [12] Reilly D, Echenique A, Wu A, Tang A, Edwards WK. Mapping out work in a mixed reality project room. In: Proc. of the 33rd Annual ACM Conf. on Human Factors in Computing Systems (CHI 2015). ACM, 2015. 887-896.
- [13] Wobbrock JO, Morris MR, Wilson AD. User-defined gestures for surface computing. In: Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (CHI 2009). ACM, 2009. 1083-1092.
- [14] Pereira A, Wachs JP, Park K, Rempel D. A user-developed 3-D hand gesture set for human-computer interaction. Human Factors, 2015,57(4):607-621.
- [15] Turunen M, Hakulinen J, Melto A, Heimonen T, Keskinen T, Hella J. SUXES-user experience evaluation method for spoken and multimodal interaction. In: Proc. of the INTERSPEECH 2009. Brighton: DBLP, 2009. 2567-2570.



黄培恺(1992-),男,陕西西安人,硕士,CCF 学生会员,主要研究领域为人机交互.



史元春(1967-),女,博士,教授,博士生导师,CCF 会士,主要研究领域为人机交互,多媒体,普适计算.



喻纯(1984-),男,博士,副教授,CCF 高级会员,主要研究领域为人机交互,图形用户界面,VR/AR 技术和应用,智能空间交互.