International Conference on Artificial Reality and Telexistence
Eurographics Symposium on Virtual Environments (2017)
R. Lindeman, G. Bruder, and D. Iwai (Editors)

Won by a Head: A Platform Comparison of Smart Object Linking in Virtual Environments

B. Ens,[1,2] F. Anderson,[1] T. Grossman,[1] M. Annett,[1] P. Irani,[2] and G. Fitzmaurice[1]

[1]Autodesk Research, Canada
[2]University of Manitoba, Canada

Abstract

Mixed -reality platforms and toolkits are now more accessible than ever, bringing a renewed interest in interactive mixed-reality applications. However, more research is required to determine which available platforms are best suited for different situated tasks. This paper presents a user study that compares headworn and handheld platforms with a smart object linking task in interactive virtual environments. These platforms both have potential benefits for supporting spatial interaction for uses situated in the spatial context of the objects being connected. Results show that the immersive, headworn platform has several benefits over the handheld tablet, including better performance and user experience. Findings also show that semantic knowledge about a spatial environment can provide advantages over abstract object identifiers.

CCS Concepts

•***Human-centered computing*** → ***Virtual reality; Interaction devices;*** *Mixed / augmented reality;*

---

1. Introduction

There is currently a growing interest in mixed-reality applications due to the recent availability of hardware platforms [HTC17] and software development tools [App]. One benefit of these mixed-reality platforms is their support spatial interaction, which allows for natural body motions. Spatial interaction methods have been shown to allow more effective interaction for some analytic tasks than traditional abstract navigation [BN08,EFI*14,LCBL*14], due to their affordance for proprioception, spatial memory and bimanual input. Spatial interaction may prove particularly advantageous for situated applications such as in-situ data visualization, virtual environment authoring, or managing connections between smart objects. However, few studies have investigated the advantages and trade-offs of available technological platforms to determine how well the benefits of spatial interaction are realized.

To better understand the trade-offs of these platforms for interactive virtual environments, we conducted a study to compare two common spatial viewing platforms: a tracked handheld tablet and an immersive headworn display (Figure 1). Handheld devices are now ubiquitous and numerous '360°' applications are available for watching video, viewing scenery or playing games. Recently available headworn platforms are more expensive and less portable than handheld devices, but are intended to provide a richer, more immersive experience.

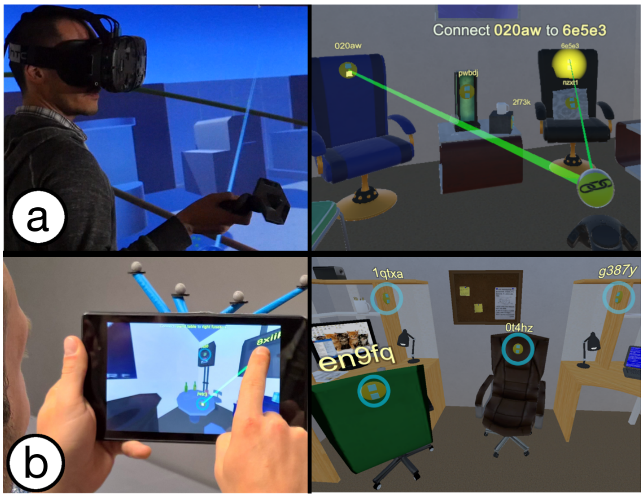

Figure 1: *A user study compares an object linking task using a headworn HTC Vive versus a handheld tablet. a) The headworn platform uses a virtual ray for object selection. b) A tracked handheld tablet uses direct touch input for object selection and. In both interfaces, a green line indicates a partially completed link.*

While there are also handheld and headworn platforms that support augmented reality (AR) for fully-situated experiences, we begin currently by looking at interactive virtual experiences, focusing on the implications of platform differences on interface design. Whereas popular 360° applications are not considered true VR, we hereafter use the term 'VR' to denote 6-dof-tracked, interactive ex-

DOI: 10.2312/egve.20171333

periences situated in virtual environments. We apply this term to both platforms to differentiate from fully-situated AR.

Our study results show that the headworn platform allows faster linking, with lower task load and greater presence. This implies that a headworn platform can provide greater advantages for spatial interaction than a handheld interface, and may be better suited for situated tasks such as object linking.

2. Object Linking Study

Our study takes place in the context of situated authoring for the Internet of Things (IoT). Researchers have made recent efforts to assist the understanding of complex smart object networks by overlaying visual information directly onto the immediate environment. Recent works have explored such situated visualizations using projected overlays [VSLC09], handheld augmented reality (HAR) [HKM13] and virtual reality (VR) [EAG*17]. Virtual environments are potentially useful in cases where the real environment is not readily accessible, for instance to simulate remote locations and dangerous environments, or to preview buildings that are still in the planning stages.

Whereas the simple task of object selection has been well-studied in virtual environments, we chose a somewhat more complex task of object 'linking' — creating a connection between a pair of spatially situated smart objects. As linking is a fundamental operation for authoring IoT programs, it must be well supported by spatial interface designers. We compare two platform implementations based on recent IoT authoring tools [EAG*17, HKM13]

The benefits of spatial interaction may partially arise from semantic knowledge about the environment, such as object descriptions and relative spatial layout. Such information is not readily available to users of traditional, desktop programming tools, who typically rely on abstract labels to differentiate objects. To investigate effects of the availability of semantic knowledge, we include a comparison of semantic descriptions versus abstract labels for identifying objects.

For evaluation, we use quantitative performance metrics (time and error) and subjective user experience measures (surveys on subjective performance and presence).

2.1. Study Task

The study task requires participants to create a directional link from one object to another. Participants are placed in a virtual environment, which they view through either a spatially tracked, *headworn* VR display or a tracked, *handheld* tablet. Among the various objects in a 5 m × 4 m room are 12 'smart' objects, denoted by yellow nodes (Figure 1). Prior to each task, participants are given instructions by display-fixed text that specifies two objects to link. Objects are denoted using either a *semantic* description of the object (e.g. large cardboard box; table lamp under painting; garbage bin by round table) or by an *abstract* label of five random alphanumeric characters (e.g. 'yr63d'). All objects within the room are hidden until the linking task begins.

To complete the task, participants must locate the two specified objects and connect a link from the first object to the second. Links that fail to join a second object are marked as incomplete, and complete links that do not join the two given objects in the correct order are marked as incorrect. After drawing a link, text feedback indicates success or failure, and the participants may proceed to the next task. Failed tasks are re-queued, with a maximum of three attempts each.

2.2. 2.2 Techniques and Platform Implementations

We created four virtual environments (two for training and two for study trials) in Unity, which are viewed using the headworn and handheld platforms. Both devices are spatially tracked to allow freedom of movement within the virtual environment, and both provide the same virtual camera field of view width (110°), so performance differences should be primarily due to properties of each platform's interface, and not confounded by viewing limitations.

Each platform has specific trade-offs. For instance, the handheld tablet provides a firm surface to facilitate precise touch input [LSH99]. The headworn interface, in contrast, uses raycasting, which also provides direct input but is known to suffer from hand jitter [ST13]. Also, whereas the handheld tablet must be always aligned for use as a viewing portal, the headworn interface also allows the head and pointing arm to move independently, for increased freedom.

For the headworn VR platform, we use an HTC Vive. The Vive provides a stereo view, with a resolution of 1080 × 1200 for each eye and a FoV of 110°. It is tracked by an external proprietary system within a 2.1 m × 2.1 m area and is tethered to a PC with a GeForce GTX 1070 graphics card, with a refresh rate of 90 fps.

The headworn interface is based on the linking tool from Ivy [EAG*17], an immersive system for authoring situated IoT programs in VR. Object selections are made using a virtual ray attached to a single controller, which is visible in the virtual environment (Figure 1a). To begin drawing a link, participants point at an object and press the trigger button. A green line connects the selected object to the controller, until the trigger is released, resulting in either a completed (correct or incorrect) link or an invalid trial.

For the handheld platform, we use a Nexus 7 tablet, with a resolution of 800 × 1280 and a weight of 340 g, running at 30 fps.

The handheld interface is modelled after the Reality Editor implementation [HKM13], an AR object-linking tool that uses marker-based tracking. In contrast to the Reality Editor's AR implementation, which allows users to see the real-world objects, our study participants can view objects only through the device portal. However, to control for viewing angle, we increase the tablet's virtual camera field to a width of 110°. For best performance and reliability, we track the tablet externally with a Vicon system. As in the Reality Editor, participants draw a link by tapping and 'dragging' a finger from one object to another, with selectable objects marked by rings (Figure 1b). Alternatively, users may hold the finger in place, while moving the tablet to align the finger with the second object.

2.3. 2.3 Study Design and Participants

The study used a within-subjects design with two factors: VR platform (headworn, handheld), and object reference (abstract, semantic). Each participant completed four blocks of trials on each platform. A block contained 12 trials, with each trial referencing one pair of randomly selected objects. Each object was included twice per block, with one trial identifying it via an abstract reference and one with a semantic reference. Presentation of tasks with abstract and semantic references was alternated to distribute any potential learning effects, and platform order was fully balanced between participants. Object locations and labels were static, so participants would benefit from spatial memory over the course of the four blocks. Each platform used a different environment to prevent carry-over of learning.

We invited 12 participants with a broad range of programming experience. All were right handed males, either students or employees of our university, and were of mean age 25 (SD 6.3). Each participant received a $25 voucher for a session lasting between 60-90 minutes.

As performance metrics, we recorded the time and success rate for each task. After using each platform, participants completed a NASA 'raw' TLX form to measure perceived task load and an ITC-SOPI [LFKD01] questionnaire to measure their sense of presence. The ITC-SOPI was designed for comparison of presence on different virtual platforms, with questions clustered into four categories: Spatial presence, engagement, ecological validity / naturalness, and negative effects.

2.4. Study Results

Errors: The number of failed trials for each platform are summarized in Figure 2a. Of a total of 1334 recorded trials, 88 (6.6%) were incomplete (i.e. did not complete a link to a second node). Because the error rates were not normally distributed, we used Wilcoxon signed ranks tests. Results showed a significant difference in the percentage of incomplete trials for both platform (headworn 3.1%, handheld 9.1%, Z=-2.67, p=0.008) and for reference (abstract 4.9%, semantic 7.4%, Z=-2.20, p=0.028).

Of the 1246 complete trials, 100 (8.8%) were marked as incorrect (i.e. one or both of the linked objects were incorrect). There was no significant difference in the number of incorrect trials for platform (headworn 7.1%, handheld 8.6%, Z=-1.17, p=0.241), but there was a difference for reference (abstract 2.9%, semantic 12.4%, Z=-2.20, p=0.028).

Task Time: Trial times are summarized in Figure 2b. For time analyses, we included only correct links (1146 trials), with outliers $\geq$ 3 SD excluded (15 trials, 1.3% of correct trials). A 2$\times$2 repeated measures ANOVA revealed a significant main effect of platform ($F1,11=164.1$, $p<0.001$, $\eta^2=0.545$). Participants were faster with the headworn (6.4 s, SD 3.9) than the handheld (10.6 s, SD 5.3) platform. There was also a significant effect of reference ($F1,11=96.0$, $p<0.001$, $\eta^2=0.389$), with semantic information (7.0 s, SD 4.5) leading to faster linking times than abstract labels (10.0 s, SD 5.2). No significant interaction effect was found ($F1,11=0.13$, $p=0.911$, $\eta^2<0.001$).

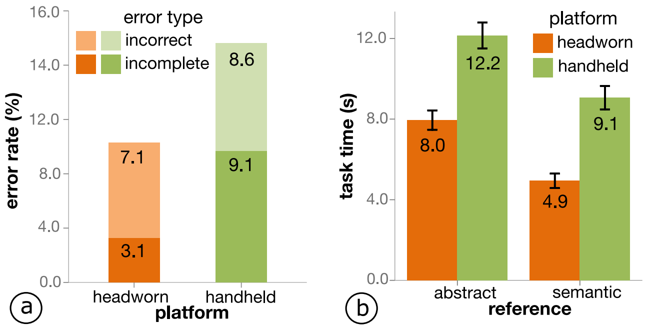

Figure 2: *a) Percentage of failed linking tasks for each platform. b) Task time by platform and reference type (bars show ±2SE).*

Questionnaires: Mean scores from the raw-TLX are shown in Figure 3a. Wilcoxon signed rank tests showed significant differences for all categories except temporal demand, with the headworn platform being lower in all cases.

Mean scores for the ITC SOPI are shown in Figure 3b. Wilcoxon signed rank tests showed significant differences in all categories except for negative effects, with the headworn platform scoring higher in the three significant categories.

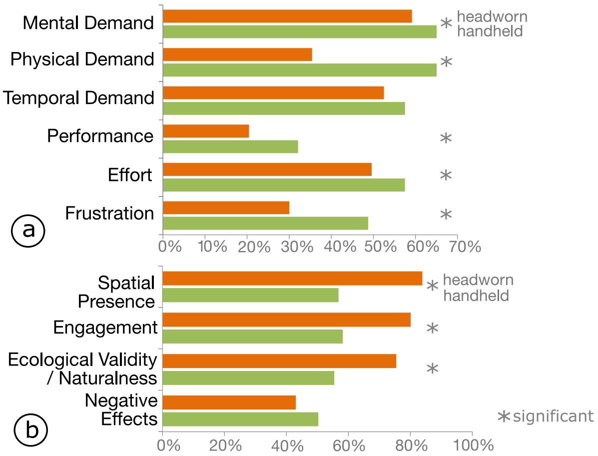

Figure 3: *a) Mean NASA raw-TLX scores (lower scores indicate lower task load) and b) Mean ITC SOPI scores by VR platform (higher scores indicate a more pronounced experience).*

3. Conclusion

A comparison of platforms for a virtual IoT linking task shows faster performance, and a better user experience for the an immersive headworn platform over a handheld tablet platform. Participants were also able to leverage semantic knowledge of the situated environment to better advantage than abstract labels on the situated objects. This work demonstrates some of the strengths of recently available wearable platforms for spatial interaction on situated tasks such as authoring of IoT links.

This study compared only virtual versions of these platforms, and may not generalize to augmented reality implementations. A handheld AR platform would provide advantages over the current 'VR' implementation. Meanwhile, current headworn AR platforms have several restrictions, such as greater weight and limited FoV, that do not apply to in the headworn VR platform used for this study. Furthermore, mobile AR workers will likely prefer smaller, lightweight input controllers than those available with home VR systems, which may negatively impact performance.

In future research, we would like to conduct a similar study using common AR platforms to determine their relative trade-offs. This future evaluation will ideally include a greater range of tasks than the simple linking task studied here. Further work will explore other applications that can benefit from situated spatial interaction, such as authoring information layouts for office and industrial settings, and situated analytics for environmental sensor data.

References

[App] Apple: Introducing ARKit. URL: https://developer.apple.com/arkit/. 1

[BN08] Ball R., North C.: The effects of peripheral vision and physical navigation on large scale visualization. *Proceedings of Graphics Interface 2008 (GI '08)* (2008), 9–16. URL: https://dl.acm.org/citation.cfm?id=1375717http://portal.acm.org/citation.cfm?id=1375714.1375717. 1

[EAG*17] Ens B., Anderson F., Grossman T., Annett M., Irani P., Fitzmaurice G.: Ivy: Exploring spatially situated visual programming for authoring and understanding entelligent environments. *Proceedings of Graphics Interface 2017 (GI '08)* (2017), 156–162. URL: https://dl.acm.org/citation.cfm?id=3141507, doi:10.20380/gi2017.20. 2

[EFI*14] Ens B. M., Finnegan R., Irani P. P., Ens B. M., Finnegan R., Irani P. P.: The personal cockpit: A spatial interface for effective task switching on head-worn displays. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems (CHI '14)* (2014), ACM Press, pp. 3171–3180. URL: http://dl.acm.org/citation.cfm?doid=2556288.2557058, doi:10.1145/2556288.2557058. 1

[HKM13] Heun V., Kasahara S., Maes P.: Smarter objects. In *CHI '13 Extended Abstracts on Human Factors in Computing Systems - CHI EA '13* (2013), ACM Press, p. 961. URL: http://dl.acm.org/citation.cfm?doid=2468356.2468528, doi:10.1145/2468356.2468528. 2

[HTC17] HTC: Vive, 2017. URL: https://www.vive.com/. 1

[LCBL*14] Liu C., Chapuis O., Beaudouin-Lafon M., Lecolinet E., Mackay W. E., Liu C., Chapuis O., Beaudouin-Lafon M., Lecolinet E., Mackay W. E.: Effects of display size and navigation type on a classification task. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems (CHI '14)* (2014), ACM Press, pp. 4147–4156. URL: http://dl.acm.org/citation.cfm?doid=2556288.2557020, doi:10.1145/2556288.2557020. 1

[LFKD01] Lessiter J., Freeman J., Keogh E., Davidoff J.: A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators and Virtual Environments 10*, 3 (jun 2001), 282–297. URL: http://www.mitpressjournals.org/doi/10.1162/105474601300343612, doi:10.1162/105474601300343612. 3

[LSH99] Lindeman R. W., Sibert J. L., Hahn J. K.: Towards usable VR: An empirical study of user interfaces for immersive virtual environments. . In *Proceedings of the SIGCHI conference on Human factors in computing systems (CHI '99)* (1999), ACM Press, pp. 64–71. URL: http://portal.acm.org/citation.cfm?doid=302979.302995, doi:10.1145/302979.302995. 2

[ST13] Stuerzlinger W., Teather R. J.: Considerations for targets in 3D pointing experiments. *Proceedings of HCI Korea (HCIK '15)* (2013), 162–168. URL: https://dl.acm.org/citation.cfm?id=2729510https://sslgate.uni-regensburg.de/,DanaInfo=dl.acm.org+citation.cfm?id=2729549{&}CFID=976130394{&}CFTOKEN=16957656http://ws.iat.sfu.ca. 2

[VSLC09] Vermeulen J., Slenders J., Luyten K., Coninx K.: I Bet You Look Good on the Wall: Making the Invisible Computer Visible. Springer, Berlin, Heidelberg, nov 2009, pp. 196–205. URL: http://link.springer.com/10.1007/978-3-642-05408-2{_}24, doi:10.1007/978-3-642-05408-2_24. 2