Touch-based Eyes-free Input for Head-Mounted Augmented Reality Displays

Tintu MATHEW, Martin WESTHOVEN, Jessica CONRADI, Thomas ALEXANDER

Department of Human Factors, Fraunhofer FKIE
Zanderstr. 5, 53177 Bonn, Germany

Abstract: Interacting with head-mounted augmented reality displays using natural user interfaces like speech recognition or gesture recognition is not practical in many situations notably in public spaces. Since these displays can be used combined with smartphones or wearables like smartwatches, the user interaction elements can be distributed across these devices. An eyes-free touch input concept for implementation on a connected mobile device with a touchscreen is presented here. An experiment was carried out to investigate the user performance on three different input devices, a smartphone, a smartwatch and a head mounted touch panel (HMT), using the same set of the touch gestures. Mobile devices are often used while walking; accordingly the interaction was investigated both while standing and walking. The evaluation showed no significant difference in user performance, response time or errors. A significant difference in subjective performance between the HMT and the smartphone was found using the NASA TLX questionnaire. As expected, the subjective estimation of mental, physical and temporal demands as well as effort and frustration were significantly higher while walking compared to standing. It was expected that the size of the touch screen would affect the performance of the various input devices, but this could not be verified. The touch gestures used were well suited for all three touch devices. Since the HMT is an integrated controller and there were no significant drawbacks in terms of performance compared to smartphone or smartwatch it is worth exploring how to optimize gesture control, size and placement of HMTs for future augmented reality displays.

Keywords: Wearables, eyes-free input, augmented reality user interfaces

1. Introduction

An easy and fast interaction technique can help increase the usability and open up more application contexts for augmented reality head-mounted displays. Commonly used input techniques for head mounted displays include natural user interfaces that use gesture, speech and gaze recognition, touch panels attached to the displays as well as handheld input devices. Natural user interfaces in particular are attracting increased attention with popular augmented reality displays like the HoloLens (Microsoft 2017) and Meta glasses (Meta Company 2017) focusing on natural user interfaces as their primary interaction method. Gesture based interfaces are often cumbersome to use and in many instances not suitable for public spaces. Robust speech recognition is problematic in noisy environments and the user’s ability to discretely interact with the display is limited. In situations like this, a smartphone or a wearable like smartwatch that the user has at hand can act as an input device and can provide private and reliable interactions with the augmented reality interface.

Touch gestures on an external input device can be employed to interact with augmented reality interfaces in head mounted displays. Ham et al. (2014) employed the touchscreen and inertial measurement unit of a smartwatch like wristband to interact with smart glasses. Benko et al. (2005) a touch table to interact with HMD in a multi-user collaborative environment. They also explored private interactions using specific gestures in a collaborative environment. Budhiraja et al. (2013) used touch gestures as well as motion of a hand held device to interact with a head mounted display. They suggest that touch gesture based input might be preferred more than device motion based gestures. Yoon et al. (2015) studies less noticeable interactions with devices like smartglasses using a finger-worn textile input device. In their pilot study, they argue that rich interactions are possible with one finger. Tsai et al. (2016) introduces a Thumbring for discrete interactions with smartglasses. The Thumbring is equipped with an inertial measurement unit to track motion and the users touch and slide finger segments to interact with content.

The user can use the touchscreen of a smartphone or other wearables as a touchpad without having to look at it, thereby avoiding visual distraction and an overlap of the input device interface and augmented reality content. They can also offer single handed interaction. Eyes-free touchscreen interaction design should reduce the need for selection accuracy and use intuitive gestural mappings to allow fast interactions with the user interface and navigation (Kane et al. 2008). Gestures which need the user to be aware of the relative location of their finger on the screen should be avoided (McGookin et al. 2008) . During single handed interaction with a smartphone, the thumb is usually used to perform touch gestures. The gestures considered for input should consider human anatomical constraints and provide an intuitive mapping from gestures to system state (Oakley & Park, 2007). Moving the thumb diagonally along devices of any size is difficult and should therefore be avoided for input tasks. Horizontal and vertical thumb movements are recommended for repetitive tasks (Karlson et al. 2008).

Not much attention has been paid to the usability of smartphones or other commercial-off-the-shelf wearables as input devices for head mounted augmented reality displays. In order to gain more detailed insights, an experiment was conducted to investigate how the attachment and holding as well as the haptics of wearables and mobile devices affect performance, distraction and subjective stress.

2. Method

The extent to which user performance in touch based interaction with head mounted display is influenced by the physical attributes and placement of or holding input devices was investigated. Three touch-based input devices are compared here, a smartphone held in the hand, a smartwatch worn on the arm and a touch panel attached to the head mounted augmented reality display. The size of the touchscreen varies across the devices. It was hypothesized that performance varies on the different input devices. Input devices held in the hand cause less stress than input devices carried on the arm or head. It was hypothesized that the head-mounted input device causes the highest stress. Since mobile devices are often used in motion and running motion can have an influence on performance, distraction and stress, this is investigated both while standing and walking. It was hypothesized that the performance of a task increases when it is solved while standing up instead of walking.

The study was carried out with 16 volunteers aged 29.4 ± 4,5 (Mean ± SD). A design with repeated measures for all conditions was chosen, conditions were permutated according to Latin square (Bortz 1993).

The participants' main task is to observe the surroundings during virtual tour of a village and to pass on certain information using the augmented reality (AR) interface. The houses or landmarks in the virtual environment is marked using a specific code. The codes consist of a number from 1 to 8 and a letter from A to H, making a total of 64 combinations. The codes on a set of houses are highlighted by a turquoise rectangle in the AR view. If a code is selected, the selection is highlighted in yellow. In addition, the label "to be edited" or "irrelevant" is now displayed using symbols as shown in Figure 1. The characters on the relevant landmark must be entered using a menu displayed in the AR view with one of the interaction devices. After completion of a task, the next set of houses is displayed.

Figure 1: Alphanumeric codes on landmarks are highlighted using a turquoise rectangle in the AR view. A selected rectangle is highlighted in yellow and the symbols for "to edit" (left) and "irrelevant" (right) is displayed on the lower right corner of this rectangle.

In addition, an attention demanding secondary task was administered to distract the user. Randomly generated distractors appearing in the AR view can be neutralized by temporarily focusing on the distractor. The distractor appears as a red blob if it is in the field of view (FOV) of the user; otherwise an arrow appears indicating the position of the distractor as shown in Figure 2. A gray or transparent blob indicates the area of focus of the participant. If the task is not fulfilled after three seconds, this will be considered an error. A successful or failed task is acknowledged with specific tones and a new distractor is displayed at randomly generated intervals. These intervals, as well as the time for neutralization, were determined in preliminary tests and set to a range of 1.5 to 6 seconds.

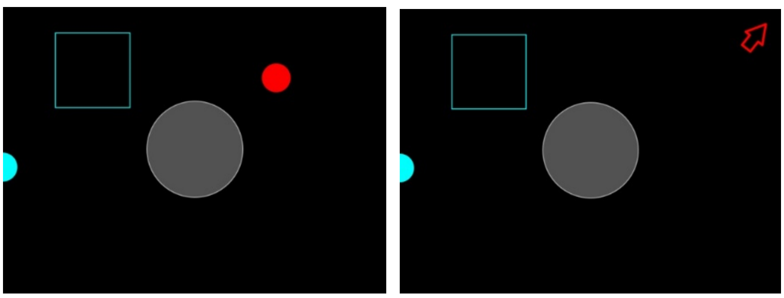

Figure 2: Visualization of the distractors in the augmented reality interface. The distractors in the user's FOV appear as a red blob, and a red arrow indicates the position of the distractors outside the FOV.

The types of user interactions required in this AR interface includes object selection and menu interaction. The user interacts with the objects in the 3D environment as well as a 2D menu using touch gestures like single tap, vertical and horizontal swipes as well as long press.

The distributions of all experimental variables were tested for normality (Kolmogorov-Smirnov test) and sphericity (Mauchly test); in case of significant results, the variance analysis uses the value corrected according to Greenhouse-Geisser. A multivariate variance analysis with repeated measurements is then performed. For dependent variables with significant differences, comparison is then performed in pairs with Bonferroni correction (IBM SPSS Statistics for Windows, 2011).

3. Results

The performance in the main task was evaluated based on the dependent variables correctness of input and superfluous input steps. For the dependent variable correctness of input, the variance analysis did not reveal any significant differences between different input devices ($F_{(3, 42)}$ = 2.782; p=0.053; η^2=0.166) or between walking and standing ($F_{(1, 14)}$ = 0.887; p=0.362; η^2=0.06). Also for the dependent variable superfluous input steps, no significant differences was observed between different input devices ($F_{(1.625; 22.752)}$ =2.782; p= 0.149; η^2=0.132) or between walking and standing ($F_{(1, 14)}$ = 1.144; p=0.303; η^2=0.076).

The performance in the secondary task was determined by the number of distractors detected and the average reaction time to the distractors. For the average reaction time to distractors, a statistical trend was evident for the variable walking. For the dependent variable number of distractors detected, the analysis did not show any significant differences between different input devices ($F_{(1.502, 21.023)}$ =2.179; p= 0.147; η^2=0.135) or between walking and standing ($F_{(1,14)}$ = 2.371; p=0.146; η^2=0.145). Also for the variable average reaction time to the distractors, no significant influence of the different input devices ($F_{(1.881, 26.330)}$ =0.897; p= 0.414; η^2=0.060) or walking and standing ($F_{(1, 14)}$ = 4.273; p=0.058; η^2=0.234) could be established.

NASA TLX was used to estimate the workload induced by the different input devices. Statistically significant influences of the variable device was found for the subscales Physical Demand ($F_{(1.545, 21.634)}$ = 4.273; p=0.036; η^2=0.234), Performance ($F_{(2.126, 29.763)}$ =3,409; p=0,044; η^2=0,196) and Effort ($F_{(3, 42)}$ = 3.355; p=0.044; η^2=0.193). Post-tests show significant differences for the scale performance between the conditions HMT and smartphone (p=0.039). Statistically significant influences of the variable walking were found for the subscales mental demand ($F_{(1,14)}$=15.404; p<0.01; η^2=0.524), physical demand ($F_{(1,14)}$=56.040; p<0,01; η^2=0.800), temporal demand ($F_{(1,14)}$ = 23.628; p<0.01; η^2=0.628), effort ($F_{(1,14)}$ = 25.091; p<0.01; η^2=0.642) and frustration ($F_{(1, 14)}$ = 7.441; p=0.016; η^2=0.347). The post-test carried out subsequently showed significant differences in mental demand between walking and standing (p<0.01), physical demand (p<0.01), effort (p<0.01) and frustration (p=0.016).

4. Conclusion

We studied the extent to which the placement and physical attributes of touch-based input devices, namely a head-mounted touch panel (HMT), a smartphone and a smart-watch, influence the performance of the user, especially when used while walking. The workload assessment based on NASA-TLX showed that there were statistically significant differences in physical demand, performance and frustration among input devices. The smartphone performed better than smartwatch and HMT. The HMT has produced higher physical demand, frustration and lower subjective performance compared to other devices. The size of the contact surface and the attachment of the HMT to the head could be one reason for this. It has been shown that the use of input devices while walking has a significant influence on the subjective assessment of total workload. Compared to the condition "standing", the mental, physical and temporal demands, effort and frustration were higher in the condition "walking".

There were no significant differences in objective performance measures between the different input devices in standing or walking conditions. The type of input devices did not influence performance in the main task. It was expected that the size of the touch screen would affect the performance of the various input devices, but this could not be proven. However, it should be noted that multi-touch gestures have been avoided due to the smaller, touch-sensitive range of smartwatch and HMT. The touch gestures used were well suited for all three touch devices. There were also no significant differences between the different input devices or for the factor standing/walking in the secondary task. There was a statistical trend for a longer processing time for the secondary task in walking than standing. Here, the movement of the body and the physical resources required to walk could have influenced the reaction time.

HMT and smartwatch do not need to be held in the hand during the input task, unlike the smartphone. For tasks that require the user to use their hands, it would be an advantage, as these devices do not have to be stowed away. Also assessing the input devices using objective performance measurements, none of the tested devices offers advantages or disadvantages over the others - whether walking or standing. Since the HMT is an integrated controller and has no significant performance drawbacks compared to smartphone or smartwatch it is worth investigating how to optimize gesture control and the size and placement of the touch panel.

5. References

Benko H, Ishak EW, Feiner S (2005) Cross-dimensional gestural interaction techniques for hybrid immersive environments. In: Proceedings of the IEEE Virtual Reality 2005, pp 209–327

Bortz J (1993) Statistik. Für Sozialwissenschaftler. Berlin, Heidelberg, New York, London, Paris, Tokyo, Hong Kong

Budhiraja R, Lee GA, Billinghurst M (2013) Using a HHD with a HMD for mobile AR interaction. In: IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp 1–6

Ham J, Hong J, Jang Y, Ko SH, Woo W (2014) Smart Wristband: Touch-and-Motion–Tracking Wearable 3D Input Device for Smart Glasses. In: Distributed, Ambient, and Pervasive Interactions, vol 8530, pp 109–118

IBM SPSS Statistics for Windows (2011). IBM Corp, Armonk, NY

Kane SK, Bigham JP, Wobbrock JO (2008) Slide rule. In: Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility, pp 73–80

Karlson A, Bederson B, Contreras-Vidal J (2008) Understanding Single Handed Use of Handheld Devices. In: Handbook of research on user interface design and evaluation for mobile technology, vol 1. Igi Global, 86–101

McGookin D, Brewster S, Jiang W (2008) Investigating touchscreen accessibility for people with visual impairments. In: Proceedings of the 5th Nordic Conference on Human-computer Interaction: Building Bridges, pp 298–307

Meta Company (2017) Meta Augmented Reality. https://www.getameta.com/. Accessed 13 November 2017

Microsoft (2017) Microsoft HoloLens. https://www.microsoft.com/de-de/hololens. Accessed 13 November 2017

Oakley I, Park J-S (2007) Designing Eyes-Free Interaction. In: Haptic and Audio Interaction Design, vol 4813. Springer Berlin Heidelberg, pp 121–132

Tsai H-R, Wu C-Y, Huang L-T, Hung Y-P (2016) ThumbRing. In: Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct - MobileHCI '16. ACM Press, New York, New York, USA, pp 791–798

Yoon SH, Huo K, Nguyen VP, Ramani K (2015) TIMMi. In: Proceedings of the Ninth International Conference on Tangible, Embedded, and Embodied Interaction - TEI '14. ACM Press, New York, USA, pp 269–272

Gesellschaft für
Arbeitswissenschaft e.V.

ARBEIT(S).WISSEN.SCHAF(F)T
Grundlage für Management & Kompetenzentwicklung

64. Kongress der
Gesellschaft für Arbeitswissenschaft

FOM Hochschule für
Oekonomie & Management gGmbH

21. – 23. Februar 2018

GfA Press

Bericht zum 64. Arbeitswissenschaftlichen Kongress vom 21. – 23. Februar 2018

FOM Hochschule für Oekonomie & Management

Herausgegeben von der Gesellschaft für Arbeitswissenschaft e.V.
Dortmund: GfA-Press, 2018
ISBN 978-3-936804-24-9

NE: Gesellschaft für Arbeitswissenschaft: Jahresdokumentation

Als Manuskript zusammengestellt. Diese Jahresdokumentation ist nur in der Geschäftsstelle erhältlich.

Schriftleitung: Matthias Jäger
im Auftrag der Gesellschaft für Arbeitswissenschaft e.V.

USB-Print:
Prof. Dr. Thomas Heupel, FOM Prorektor Forschung, thomas.heupel@fom.de

Screen design und Umsetzung
© 2018 fröse multimedia, Frank Fröse
office@internetkundenservice.de · www.internetkundenservice.de