A Dose of Reality: Overcoming Usability Challenges in VR Head-Mounted Displays

Mark McGill[*]
m.mcgill.1@research.gla.ac.uk

Daniel Boland[*]
daniel@dcs.gla.ac.uk

Roderick Murray-Smith
roderick.murray-smith@glasgow.ac.uk

Stephen Brewster
stephen.brewster@glasgow.ac.uk

School of Computing Science
University of Glasgow, UK

ABSTRACT

We identify usability challenges facing consumers adopting Virtual Reality (VR) head-mounted displays (HMDs) in a survey of 108 VR HMD users. Users reported significant issues in interacting with, and being aware of their real-world context when using a HMD. Building upon existing work on blending real and virtual environments, we performed three design studies to address these usability concerns. In a typing study, we show that augmenting VR with a view of reality significantly corrected the performance impairment of typing in VR. We then investigated how much reality should be incorporated and when, so as to preserve users' sense of presence in VR. For interaction with objects and peripherals, we found that selectively presenting reality as users engaged with it was optimal in terms of performance and users' sense of presence. Finally, we investigated how this selective, engagement-dependent approach could be applied in social environments, to support the user's awareness of the proximity and presence of others.

Author Keywords

Virtual Reality; Augmented Virtuality; Engagement;

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

Virtual Reality (VR) is seeing a resurgence as a medium for work and entertainment [23]. New consumer head-mounted displays (HMDs) take advantage of low weight, low cost, high resolution displays, delivering rich and immersive VR experiences. This is VR's greatest strength and also it's most significant weakness; when wearing a HMD, one's visual (and often auditory) connection to the "outside" world is diminished, promoting strong feelings of presence in the virtual environment. However, even tasks as simple as picking up a cup become difficult without visual reference.

CHI 2015, April 18–23 20152015, Seoul, Republic of Korea

ACM 978-1-4503-3145-6/15/04...$15.00
http://dx.doi.org/10.1145/2702123.2702382

HMDs have recently attracted interest and investment from major companies, such as Facebook's acquisition of Oculus Rift VR, as well as Sony and Samsung's own efforts. While VR HMDs in and of themselves are certainly not new, their increasingly widespread deployment does emphasise the need to consider the user experience of consumers in home or office usage. Our work identifies the usability challenges facing consumers adopting these VR HMDs.

We take the existing body of work on VR HMDs and consider the usability of VR HMD interaction in terms of the present consumer adoption. Through a survey of current consumer HMD users, we identify a significant need to improve ways to interact with objects and peripherals, and be aware of others in the real-world environment. While virtual experiences are generally intended to be decoupled from the real environment, some interaction with real objects is nevertheless often necessary. Users reported being unable to use their keyboard or take a drink etc. when required – having to remove their HMD in order to do so. We validate the need for blending reality into virtuality in a controlled typing study, restoring some of the performance lost when typing in VR.

We then investigate ways in which we can incorporate aspects of reality into VR, firstly in terms of facilitating interaction with objects and peripherals, and then in terms of awareness of proximity and presence of others. We highlight mixed reality and transitional interfaces as a step towards addressing these issues, and link them to measures of presence and recent work on user engagement. For interaction, we examine both how much reality should be incorporated and when, finding that selectively and seamlessly blending the necessary aspects of reality based on user-engagement is optimal to preserve a sense of presence in VR. With respect to awareness, we examine the potential for applying this engagement-dependent Augmented Virtuality (AV) approach to the presence and proximity of others, finding that, while presenting awareness information visually is appropriate, there are other user-elicited ways which might be less disruptive.

This work addresses the usability issues reported by users, with a novel approach to partially blending reality and virtuality. We present *engagement-dependent* AV, selectively incorporating aspects of reality only when needed, preserving immersion and avoiding the frustration of removing a HMD in order to type, take a drink or interact with others.

---

[*] These authors contributed equally to this work.

BACKGROUND

Despite significant research and development in the 1990s, the immersive VR experiences envisaged did not reach consumers. [9] posited a number of reasons for this e.g. the technical quality of HMDs was considered poor (in terms of resolution, Field of View (FOV), comfort, motion sickness, etc.), socialization was not facilitated (with users unable to interact with others), the graphical quality of the rendered scenes was poor, and the cost was prohibitive. However many of these issues are on the verge of being addressed, e.g. the Oculus Rift's DK2 features a 1920x1080 low latency OLED display. As such, high quality HMD with accurate head tracking are on the verge of being widely available. However interacting with reality remains a challenge: existing consumer HMD such as the Oculus Rift or Gear VR[1] do not yet incorporate the sensors needed to adequately track hands, identify objects, or provide a wide-angle FOV of reality. However the rapid development of add-ons such as the Leap Motion VR[2] suggest that in the near future these HMD may have the capability to sense reality.

Interaction

Thus far, the control over an interactive system when wearing an HMD has predominantly been through gestural interfaces (e.g. motion controllers, Leap Motion) or interfaces suitable for use without sight (e.g. tangibles, handheld controllers, or relying on a subset of keyboard/mouse commands) [3]. These solutions, while potentially adequate for simple purposes, are almost certainly inadequate for any task that requires a greater bandwidth of input e.g. text entry, drawing on a pen display etc. In particular, interaction breaks down where outputs are not incorporated into the VR presentation, e.g. trying to use a phone while wearing a HMD. In the near future these issues are likely to become commonplace, with people using HMDs at their desks, or in the living room. Keyboards are a ubiquitous input technique. Their use involves both kinaesthetic feedback from fingers striking keys and visual feedback to observe hand and key placement. Virtual keyboards offer a replacement input modality in virtuality, but lack haptic feedback. This can have a dramatic effect on typing performance. In a study of typing performance on a flat keyboard (lacking the haptic feedback of key travel) versus a standard keyboard in VR, Barrett & Krueger demonstrated a significant performance drop without haptic feedback [1]. A variety of text entry methods for use when wearing VR HMD (both mobile and static) have been proposed. Outside of speech, none have been shown as approaching the performance of the standard PC keyboard [7]. There are many contexts in which speech may not be appropriate, we thus investigated how access to peripherals such as keyboards could be supported in VR.

Sense of Presence

Immersion in VR is typically quantified through the user's sense of presence. Presence can be affected by the rendering quality of the scene, the quality of the HMD's head tracking, and even how users interact with virtual objects. It can be increased through natural interactions with objects as they occur in the real world [10]. Presence can be measured in a multitude of ways, for example through brain activity, physiological measures or more traditional qualitative measures [24]. Many of these measures involve application-specific questions. We used the Igroup Presence Questionnaire (IPQ)[18] to test our designs as it is generalised and widely used in the literature. Our aim in this work can be described in terms of Slater's conception of presence as *place and plausibility illusion* [19], where we selectively blend reality only as much as needed, so as to minimise the impact on place illusion.

Blended / Mixed Reality

The field of "mixed reality" refers to displays that inhabit a point between reality and virtual reality in the "Virtuality Continuum" [15, 14]. This continuum led to the definition of both Augmented Virtuality (AV), where a virtuality view is augmented with elements of reality, and Augmented Reality (AR), where a reality view is augmented with elements of virtuality. Depending on the amount of reality or virtuality that is incorporated, a display can inhabit a very different point on the continuum. For example, a minor augmentation to reality would be close to reality on the scale, while a major augmentation to reality would be nearer to virtuality.

Augmenting the Real or the Virtual

AR has seen much research in recent years, building upon seminal work such as Navicam [17], with reality being augmented in ways that allow for novel interactive systems. In contrast, the concept of AV has received much less attention, partially due to the lack of good consumer HMDs. AV has been accomplished through the usage of chroma-key approaches toward interleaving real-world elements into a VR space. Early work in this area by Metzger [13] proposed a seamless integration of real-world human interfaces with the virtual world, using a HMD with a head-mounted camera and chroma-key image segmentation to enable video see-through. He proposed a small inset view of reality within the virtual world for a permanent view of reality, as well as the potential for making this view transparent, so that users could see through the virtual world image to a real world image.

Head-mounted cameras have been frequently used to capture reality e.g. Steinicke et al. [21] chroma-keying to the user's body, presenting a virtual hands and body in an egocentric view of virtuality. More recently, head-mounted depth cameras have allowed for hand tracking such that virtual representations of hands or objects are now becoming feasible [22], while room-wide sensors such as the Microsoft Kinect allow for user tracking, gestures and physiological measures, all of which could potentially be used to augment virtuality with information about reality that cannot necessarily be captured from a head-mounted camera. There remains the question of managing this augmentation of virtuality i.e. managing traversals or transitions within the virtuality continuum as discussed by Davis et al. [6]. Most research has examined mixed reality boundaries [2] where physical boundaries mark crossovers between mixed and virtual reality [8]. Where an interaction spans distinct points on the mixed reality continuum, it is termed a *transitional interface*. The importance of continuity in transitional interfaces has been highlighted, as well as the lack of a theoretical guideline for when to make these transitions [5], and how much to transition by.

---

[1] `http://www.samsung.com/global/microsite/gearvr/`

[2] `https://www.leapmotion.com/product/vr`

LINKING ENGAGEMENT TO MIXED REALITY

Pohl & Murray-Smith's focused–casual continuum describes interaction techniques according to the degree to which they allow users to adapt how much attention and effort they choose to invest in an interaction, conditioned on their current situation, i.e. the ability to adapt how engaged they are [16]. They take a control-theoretic approach to defining engagement; the more engaged with an interaction users are, the more frequently and accurately they sample feedback and provide more rapid and accurate high-bandwidth input.

Depending on their current context in VR, users will vary in how much they wish to engage with interactive objects in reality. Users may wish to incorporate real persons around them or interactive devices like keyboards or phones into their VR experiences e.g. [4] used chroma-keying to incorporate real-world tools into VR according to architectural context when exploring a VR house. We suggest that transitions in mixed reality can be linked to user engagement with interactive elements in the different environments. A user exploring a VR environment may have no need for objects in reality. However, if they extend their hands to engage with their keyboard, there should be a transition to a mixed reality mode. Allowing users to choose a level of engagement and control their position on the mixed reality continuum in this natural way combines transitional interfaces, mixed reality and focused-casual control.

We adopt the approach taken in the focused-casual work to show how engagement-dependent AV supports richer interactions with higher bandwidth input. In Figure 1, we show the control loops involved in our mixed reality system (blue for real, red for VR). The VR control loop involves input from reality and feedback from the virtual environment. This control loop is high-level, comparable with on-screen feedback for typing: users see the results of their typing but not their hands. The more a user engages with a real interactive object such as a person or keyboard, the more we can blend it into the VR view. Users can thus perceive the low-level feedback from interacting with the real object, e.g. hitting keys when typing or social signals from proximate persons. Their control loop with reality is thus more tightly-coupled and they can rapidly sample detailed feedback from reality.

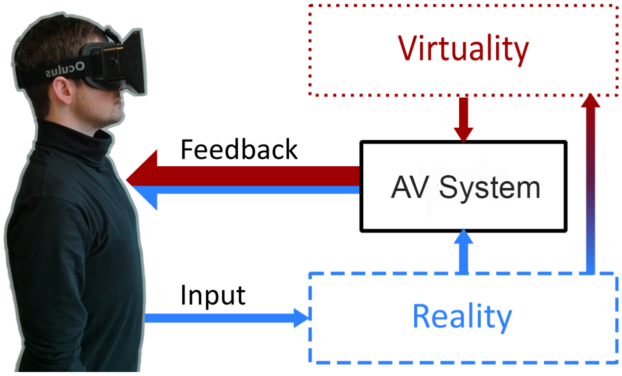

Figure 1. Feedback from real objects being engaged with can be used to augment virtuality, providing a low-level control loop (in blue) that can support high bandwidth interaction, such as typing. The more the user engages with reality, the more real feedback is mixed with virtuality.

VR HMD USABILITY SURVEY

We developed a survey to elicit general usability concerns, to examine how users interact with peripherals and objects, and impediments to VR use. A second, more focused, stage of the survey investigated the extent to which interaction with, and awareness of, reality posed significant concerns for the usability of VR HMDs. The survey was sent to mailing lists (covering University staff, students and HCI practitioners), as well as online forums and VR-related communities, receiving 108 responses in total. Questions were not forced choice.

Existing Usage

The most used headsets were the Oculus Rift DK1 (used by 49% of respondents) and DK2 (79%), with others such as Google Cardboard having been used by under 5% of respondents. 75% of respondents used VR HMDs weekly or more frequently, with 36% using them daily. In terms of auditory feedback, headphones were used by the majority of respondents (82%: In-Ear (13%), Enclosed (57%), Noise Cancelling (10%), and Bone-conduction (1%)), with a minority (17%) using speakers. Of interest is the fact that 80% of users actively try to cancel out any perception of real-world audio.

Impediments To VR Usage And Enjoyment

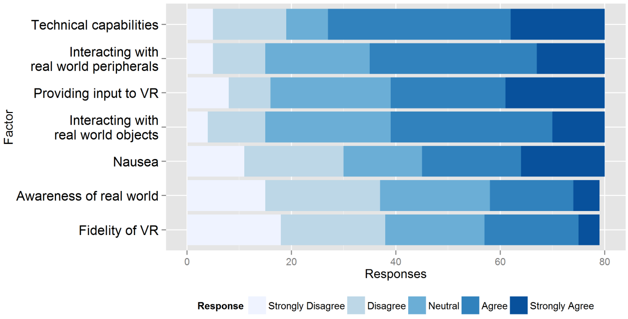

Figure 2. "To what extent do you agree that the following IMPEDE your ability to use and enjoy VR HMDs?". Aspects ordered by mean score highest to lowest, areas questioned were: *Technical capabilities of headset* (e.g. head tracking, resolution, latency); *Nausea when using headset*; *Fidelity of virtual world* (e.g. how real does it look); *Awareness of real world* (e.g. who is there); *Interacting with real world objects* (e.g. picking up a cup); *Interacting with real world peripherals* (e.g. via keyboard, mouse, motion controllers..); *Providing input to the virtual world* (e.g. via peripherals, gesture, voice etc.); n=81.

We asked about potential impediments to the usage and enjoyment of HMDs to ascertain the relative importance of interaction with reality versus typical complaints regarding VR HMDs (e.g. in terms of nausea, poor resolution, etc.). While the technical capabilities of HMDs dominated, interaction with peripherals and real world objects were rated highly, with over half of respondents agreeing or strongly agreeing that these were currently impediments to their HMD usage.

Interacting With Reality

To interact accurately with the real world, users currently need to remove the HMD. We first sought to establish how often they interrupt their VR experience to do so, and how frustrating they find this. We found that, while this is not a particularly frequent occurrence, it is frustrating for users to have to resort to this behaviour. Figure 3 shows the distribution of responses for frequency with respect to taking the headset off versus frustration.

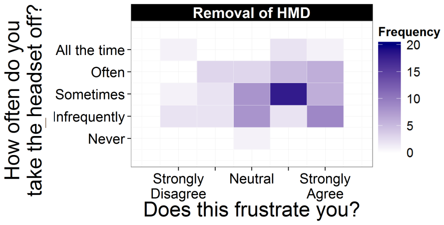

Figure 3. Responses to "When using the VR HMD, how often do you take the headset off (lifting it off your eyes temporarily, removing it etc.) in order to interact with, or gain awareness of, the real world?" and "Having to remove the VR HMD in order to interact with reality frustrates me". Colour indicates frequency, darker is more frequent. n=81.

We then asked users to gauge how effectively they currently manage to interact with three aspects of reality: objects, others (proximate persons) and peripherals, as well as to what extent they agreed that VR HMDs should better facilitate interaction with these, as seen in Figure 4.

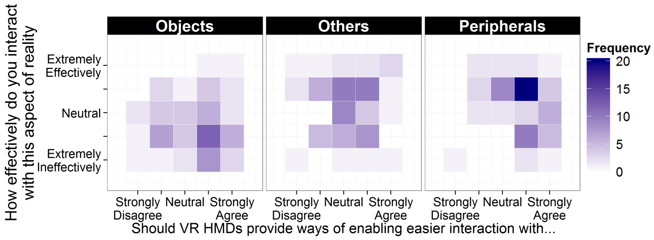

Figure 4. Responses to "How effectively do you interact with the real world when wearing a VR HMD in the following contexts" and "Do you agree that VR HMDs should provide ways of enabling easier interaction (e.g. seeing / hearing) with the following". n=76.

Interaction with 'objects' is mostly rated extremely ineffective, with strong agreement that such capability should be better facilitated. Interacting with peripherals is somewhat dichotomous in terms of effectiveness, with users congregating around either effectiveness or ineffectiveness with respect to their capacity to interact with others. However, the majority wish for this capability to be facilitated, bringing into question how well users can accurately gauge their capability to interact with commonly used peripherals. With respect to interacting with others, the majority considered themselves able to interact with others effectively and were neutral tending toward agree with respect to whether this should be better facilitated. This is likely strongly tied to the usage contexts: the majority of respondents polled who used VR headsets, used them in private spaces / predominantly alone (80%).

Awareness Of Environment And Others

We investigated which aspects of awareness of others and the environment were most problematic for HMD users (Figure 5). Of particular note were awareness of *Others* and their *Proximity* which were both rated poorly in terms of existing awareness, and strongly in terms of whether VR HMDs should support these aspects of awareness. Interestingly, interaction with others was distributed between unaware and aware, yet the majority of respondents agreed that this aspect of reality could be improved. *Context* was predominantly neutral on both counts, whilst explicitly *Hearing* what is happening showed a high degree of dispersal; this is presumably because of the different volume levels regarding VR and headphone types employed.

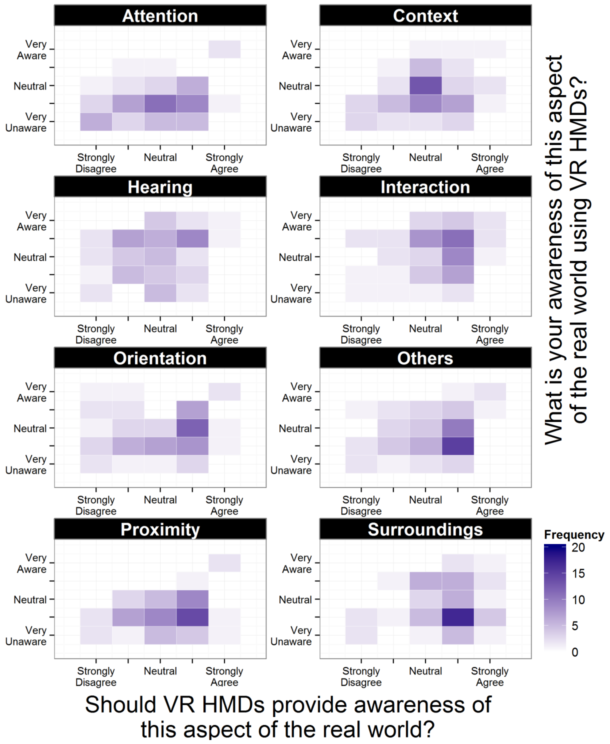

Figure 5. Aspects of awareness surveyed: *Attention* (e.g. knowing if you are being observed); *Context* (e.g. knowing what is happening around you); *Hearing* what is happening (e.g. movement, talking) *Interaction* (e.g. knowing someone is trying to talk to you); *Orientation* (e.g. knowing spatially where you are and where you are facing); *Others* (e.g. knowing others have entered the room); *Proximity* (e.g. knowing how close others are); *Surroundings* (e.g. knowing you are close to a wall). n=68.

We considered the impact lack of awareness might have on comfort and anxiety. When asked if they agreed with the statement "I experience anxiety regarding my inability to tell what is happening around me", only 22% of respondents agreed. When asked if they agreed with the statement "My level of awareness is enough that I feel comfortable and secure in my personal surroundings" only 12% disagreed. Where users lack awareness of what is going on around them, this does not appear to lead to undue anxiety or discomfort.

Survey Discussion

The survey identified users' desire for easier interaction with objects and peripherals. The (in)ability to provide input into VR was rated as a greater impediment to VR usage and enjoyment than nausea, though we note the selection bias of most respondents being VR users. We also found that, while awareness of the real world in general is not presently an impediment to usage, there are aspects of awareness that should be prioritised – the presence of others and their proximity. These issues led us to design experiments to explore potential solutions to incorporating these aspects of reality into VR.

STUDY 1: TYPING

To interact with a VR environment, a user requires an appropriate input modality. Input actions occur in reality and in VR the user loses the inherent visual feedback for these actions. Many text entry methods for VR have been discussed [7], with voice input offering reasonable performance, however adoption of novel modalities will be slow and the use of keyboards is likely to continue, even if only transitionally. The keyboard is a ubiquitous input device and widely used by gamers (currently the key demographic for consumer VR HMDs); it is familiar, its layout is consistent, it has guidance for blind typing, and users are proficient in its usage. The use of a keyboard in VR presents the immediate problem of requiring the user to locate it in reality while immersed in VR. Users must switch from considering their virtual environment to the spatial layout of their real surroundings.

Typing offers an interesting case for examining our ability to interact with reality, being an example of a rich interaction with an object that requires a high-bandwidth feedback loop. Enabling high performance keyboard use is a quantifiable means of demonstrating the general opportunity for rich interaction supported by engagement-dependent AV. As such, we resolved to bring the real-world keyboard into virtuality, hypothesising that this would demonstrate an increase in performance of blending feedback from reality with virtuality.

Design

We conducted a text entry study with the following conditions (depicted in Figure 6): *(1) Reality:* Baseline typing performance on a keyboard in full view in reality; *(2) Virtuality*: No keyboard view and wearing an HMD; *(3) Augmented Virtuality - Inferred Partial Blending*: A view of the keyboard and user's hands was blended into virtuality. *(4) Augmented Virtuality - Inferred Full Blending*: The full view of reality appeared. These conditions allowed us to assess both performance with no view of reality and the potential for improving performance through incorporating either a full view of reality, or a selected subset of it. For the AV conditions, the display of reality was inferred based on user engagement e.g. when the user reached out to use the keyboard, with the keyboard in view, that aspect of reality would be blended in.

Sixteen participants were recruited from University mailing lists (Age mean=25.6, SD=4.0, 14 male, 2 female) in a within-subjects design. For the baseline condition, users were seated at a standard single monitor workspace, while in virtuality users found themselves in a VR space with the same view as was on the monitor previously. Users were asked to enter 15 phrases, with 1 training phrase at the start of each condition to familiarise them with the keyboard view presented. Before each phrase participants were asked to put their hands by their side, in order to mimic aperiodic interaction with the peripheral from VR. Based on [11], we used the MacKenzie 500 phrase set, with phrases chosen at random and presented at the top of the display, with the bottom of the display used for on-screen feedback of typing. Baseline typing was captured first, the remaining conditions were counterbalanced. In addition to standard typing metrics [20] we measured the accuracy and time to first key press, to gauge how effectively participants could locate the keyboard.

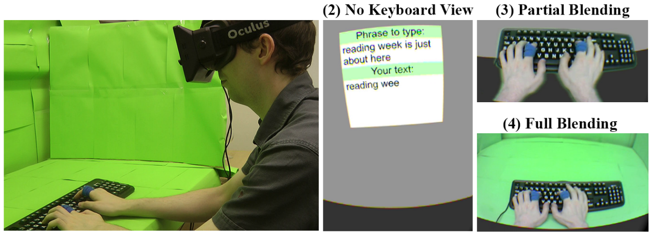

Figure 6. Experimental setup (left) and VR typing Conditions (2, 3, 4).

Implementation

To bring the keyboard into VR, we used a modified Logitech C310 HD webcam with a 1.8mm M12 wide-angle board lens mounted on the front of an Oculus Rift DK1. We chose a single camera setup with only monocular depth cues as this approximates camera setups in VR headset/phone hybrids (such as the Samsung Gear VR), as well as minimizing processor load and latency. The scene was rendered at 60FPS.

For selectively blending reality, a chroma-key approach (Figure 6) was utilized whereby the keyboard was placed within a green screen environment, with multi-threaded image processing and HSV thresholding in EmguCV used to determine the green screen contour and non-green contours within it, which were then presented in virtuality. This was performed separate to the rendering of the webcam, to minimize the latency of the user's view of reality, through an alpha mask which was generated and applied with ~1 frame of latency. Hand detection occurred within the same process with users wearing discrete markers on their hands which were detected via EmguCV blob detection. This was implemented in the Unity 4 engine (for code excerpts, see: `github.com/mark-mcg/CHI2015-VR`). The VR scene was rendered on a Intel Core i5 (3.30GHz) with an Nvidia GTX 460 GPU. The keyboard used was a standard PC size and layout, however it featured larger lettering; this was to compensate for the low resolution of the Oculus Rift HMD.

Results

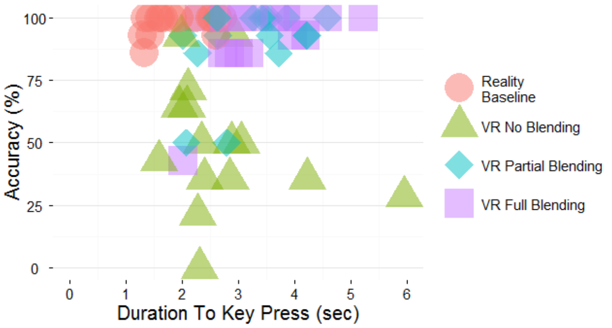

Figure 8. Mean accuracy of first key press (whether it was the correct key or not) against time to first key press, by condition and participant.

The results can be seen in Table 1 and Figure 8; a repeated-measures ANOVA (GLM) was performed with *post hoc* pairwise Tukey's tests ($\# - \#$ indicates significant difference between numbered conditions, $p < 0.05$). Baseline (1) performance was highest in most metrics, in keeping with previous work in the literature. The *status quo* (2) had significantly

Typing Metric	1 Reality Baseline	2 VR No Blending	3 VR Partial Blending	4 VR Full Blending	RM-Anova	Tukey *Post-hoc*
WPM	58.9 (17.0)	23.6 (22.3)	38.5 (19.3)	36.6 (19.7)	$\chi^2(3) = 85.1, p < 0.01$	1-2, 1-3 , 1-4, 2-3, 2-4
Total Error Rate	4.64 (3.08)	30.86 (15.17)	9.20 (6.42)	10.41 (7.36)	$\chi^2(3) = 54.19, p < 0.01$	1-2, 2-3, 2-4
Corrected Error Rate	3.77 (2.52)	30.26 (15.37)	8.13 (6.03)	8.63 (7.37)	$\chi^2(3) = 55.16, p < 0.01$	1-2, 2-3, 2-4
Not Corrected Error Rate	0.87 (1.55)	0.61 (1.37)	1.07 (1.41)	1.78 (3.52)	$\chi^2(3) = 4.17, p = 0.24$	NA
Duration To First Key (sec)	1.9 (0.5)	3.1 (1.9)	3.3 (0.8)	3.4 (0.9)	$\chi^2(3) = 22.7, p < 0.01$	1-2, 1-3, 1-4
First Key Accuracy	0.98 (0.04)	0.49 (0.24)	0.90 (0.16)	0.92 (0.14)	$\chi^2(3) = 57.7, p < 0.01$	1-2, 2-3, 2-4

Table 1. Typing statistics. Total Error Rate from [20]. First Key Accuracy is between 1 (100% accurate) and 0 (0% accurate). Green denotes $p < 0.05$

higher error rates (corrected and not corrected) and a large drop in typing rate (WPM). The two AV conditions significantly reduced the error rate in VR typing performance, nearing the baseline rate. Typing rate however, whilE improved, failed to reach baseline performance. There was no significant difference between partial and full blending.

Discussion

Augmenting virtuality with the real keyboard greatly reduced error rates as users were able to orient themselves to the keyboard as if it were in VR: users do not need to attempt to remember where the keyboard is in the real environment whilst in VR. Typing rate is a low-level feedback loop and not supported by the usual feedback in VR. Bringing the visual feedback of the real keyboard into VR helps, however a lack of stereoscopic depth cues and the latency of the VR HMD likely contributed to the modest gains in WPM in the AV conditions. These results demonstrate that incorporating reality into VR is necessary to preserve performance, and that only providing a view of the keyboard and hands does not negatively impact performance versus a full view of reality. That a partial blending of reality, as and when needed by user engagement, also enables rich interaction is a key result – validating the utility of engagement-dependent AV.

STUDY 2: INTERACTION WITH REALITY

The previous study showed that incorporating reality into VR could help interaction. This generated two further questions: how much reality should be incorporated, and when? We chose to tackle these questions within the most pressing user-elicited aspect of interaction with reality: interacting with objects and peripherals. The aim of this study was to facilitate interaction with real objects and peripherals when using a HMD while controlling the factors of *amount of blending of reality* and *control of incorporating reality*. In this way, users could experience what it was like to interact realistically using different amounts of blending and different mechanisms for incorporating reality.

Design

For this study we employed a 3×2 factorial design, with three levels for the **Blending** factor (Figure 7): $\{$*Minimal:* Including just the reality around the user's hands into the VR scene; *Partial:* Viewing all interactive objects from reality in the VR scene; *Full:* Full view of reality$\}$ against two levels for the **Control** factor: $\{$*User:* Where the user explicitly controlled the presence of reality via pressing any button on a gamepad; *Inferred:* Where the presence of reality was inferred based on user engagement, namely when their hands were in view of the camera$\}$. Additionally, we included two conditions for comparative baselines, where there was no view of reality: $\{$*Always On:* HMD was kept on at all times; *Lift Headset:* Where users were allowed to lift the HMD up in order to peek at reality$\}$. These baselines were omitted from statistical analysis as we are no longer asking the question should reality be incorporated into virtuality, but instead *how* should it be incorporated. Sixteen participants were recruited from University mailing lists (Age mean=25.3, SD=3.78, 14 male 2 female). The study used a within-subjects design. We measured the impact our factors had on users in terms of sense of presence (using the 'Igroup Presence Questionnaire' (IPQ)), and workload (NASA TLX). Additionally we solicited user rankings of preference.

Implementation

The same setup was used as in the previous study. To provide an immersive VR experience in which users could develop a sense of presence, we used a modified version of the Tuscany Villa[3] – a scene developed by Oculus to showcase the immersive capability of VR. Users were instructed on how to navigate in the scene, then given 2 minutes per condition to explore the villa as they desired using a gamepad whilst wearing the Oculus Rift DK1 and in-ear headphones. Conditions were counterbalanced.

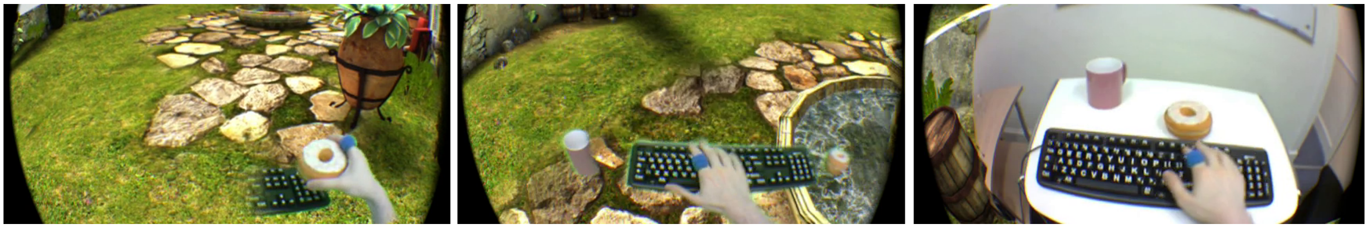

Figure 7. Left: *Minimal blending* (reality around user's hands). Middle: *Partial blending* (all interactive objects). *Right*: Full blending (all of reality)

---

[3] `share.oculusvr.com/app/oculus-tuscany-demo`

During each condition, users were prompted textually three times to interact with one of three real objects: a *keyboard* (action: type on it), a *coffee mug* (pick it up as if taking a drink), and a *toy doughnut* (pick it up as if taking a bite). These were intended to represent typical objects users might interact with when wearing an HMD. For the minimal/partial reality conditions, participants used a green-screen desk. For the full reality conditions, they used a normal desk to prevent confounding visual distraction from the green screen.

Results

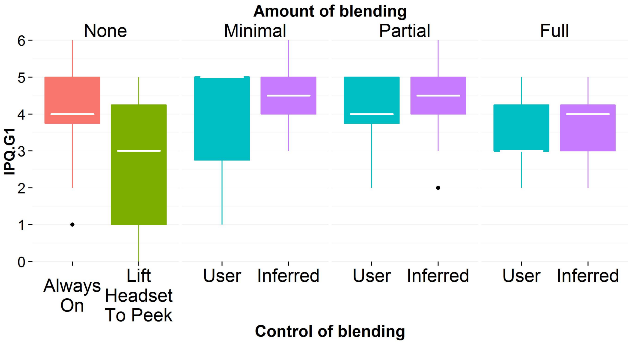

Figure 9. "Sense of being there" questionnaire (higher is better) plot of quartiles (25th, 50th, 75th) for Control and Blending factors, with comparative baselines of no blending included.

To analyse the results, a factorial repeated-measures ANOVA (GLM) was performed with contrasts where applicable. In terms of the **Control** factor (see Table 2), there were significant effects on workload, spatial presence, sense of being there, and user ranking, with inferred engagement proving superior to user control in each case. With respect to the **Blending** factor, there were significant effects on spatial presence, involvement, sense of being there, and ranking.

Contrasts revealed that, in each case, selective reality (that is, partial or minimal blending) proved superior to full reality (full blending), however there were no significant differences between partial and minimal blending. These effects are shown in Figure 9 (sense of being there) and Figure 10 (rankings). There was no interaction effect between control × blending. Modest but significant correlations were found between ranking and spatial presence ($r_s = -0.20$, $p < 0.05$), workload and realism ($r_s = -0.27, p < 0.01$), and amount of blending and spatial presence ($r_s = -0.21$, $p < 0.05$), involvement ($r_s = -0.32$, $p < 0.01$), sense of being there ($r_s = -0.22$, $p < 0.05$), and ranking ($r_s = 0.29$, $p < 0.01$).

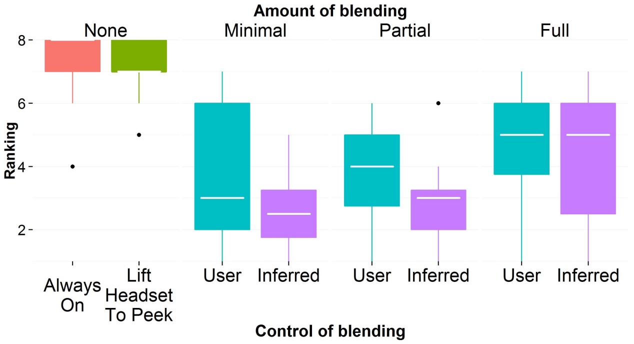

Figure 10. User ranking of preference (lower is better) plot of quartiles (25th, 50th, 75th) for Control and Blending factors, with comparative baselines of no blending included.

Discussion

Firstly, our two comparative baseline measures reaffirm that a view of reality is a necessity, with user rankings in Figure 10 rated far lower than any of the blended reality conditions. Secondly, this study confirms that inferred engagement is superior to user control of engagement. Inferring the presence of reality lowers user workload, is preferred by users and improves spatial presence and sense of being there. While inferring engagement will not work in all circumstances, this result shows that user control of blending reality should be a last resort, and that giving control to the system is preferable.

With respect to the amount of reality blended, the results clearly indicate that in terms of spatial presence, involvement, sense of being there and ranking, the full view of reality was inferior to a selective subset of reality, while there were no significant differences between partial and minimal reality. This demonstrates that in incorporating reality into virtuality we should strive to limit the amount of reality to only what is necessary to enable the capability for interaction.

Measure	Control Factor	Blending Factor	Control × Blending	Full vs. Selective Reality Contrast	Partial vs. Minimal Reality Contrast
TLX Overall Workload	$\chi^2(1) = 5.86$, $p < 0.05$	$\chi^2(2) = 4.51$, $p = 0.10$	$\chi^2(2) = 2.01$, $p = 0.37$	NA	NA
IPQ Spatial Presence	$\chi^2(1) = 4.47$, $p < 0.05$	$\chi^2(2) = 16.22$, $p < 0.01$	$\chi^2(2) = 0.29$, $p = 0.86$	$b = 0.16, t(60) = 4.11$, $p < 0.01$	$b = 0.01, t(60) = 0.14$, $p = 0.89$
IPQ Involvement	$\chi^2(1) = 0.38$, $p = 0.54$	$\chi^2(2) = 19.23$, $p < 0.01$	$\chi^2(2) = 1.13$, $p = 0.57$	$b = 0.26, t(60) = 4.49$, $p < 0.01$	$b = 0.07, t(60) = 0.71$, $p = 0.48$
IPQ Realism	$\chi^2(1) = 1.70$, $p = 0.19$	$\chi^2(2) = 3.40$, $p = 0.18$	$\chi^2(2) = 0.34$, $p = 0.84$	NA	NA
IPQ Sense of Being There	$\chi^2(1) = 4.38$, $p < 0.05$	$\chi^2(2) = 9.02$, $p < 0.05$	$\chi^2(2) = 1.22$, $p = 0.54$	$b = 0.18, t(60) = 3.01$, $p < 0.01$	$b = -0.02, t(60) = -0.15$, $p = 0.88$
Ranking	$\chi^2(1) = 4.07$, $p < 0.05$	$\chi^2(2) = 10.95$, $p < 0.01$	$\chi^2(2) = 0.78$, $p = 0.68$	$b = -0.42, t(60) = -3.30$, $p < 0.01$	$b = -0.06, t(60) = -0.29$, $p = 0.77$

Table 2. Results of two-way repeated measures ANOVA for factors Control and Blending, Green cells denote $p < 0.05$. Contrasts were performed between Full vs. Selective (Partial and Minimal) blending, and between Partial and Minimal blending.

Lifting the headset to peek at reality had a negative effect on the sense of presence. Together, these findings lead us to recommend that augmenting virtuality selectively with relevant aspects of reality based on inferred user engagement offers a solid basis for incorporating reality into virtuality whilst minimizing the effect on immersion.

STUDY 3: EXISTENCE OF OTHERS

From the results of our survey and the previous study on interaction with reality, another question arises: can we apply engagement-dependence to the most important aspects of awareness from the survey, namely the existence and proximity of others. However, within this context, engagement is significantly more problematic to define. Whereas previously, the user implicitly knew of the existence of objects to engage with and thus engagement was essentially binary, s/he may not know that there is anyone nearby in reality to engage with. Equally, those in reality can choose to engage with the HMD user, likely governed by a variety of social cues, however the user might well remain oblivious to this.

The concept of engagement here is no longer binary: there is both a necessity for a low (or casual) engagement state for a base awareness that there is potentially someone to engage with, and an accommodation for intermediary states between casual and full engagement, e.g. being aware of someone, then introducing more detail as the user engages (who is there, what are they doing, etc.). Given this, we investigated not how engagement is defined, measured and inferred in this context, but how we might communicate reality in both a low engagement state and a fully engaged state, again with an emphasis on how to minimise the impact this view of reality had on the VR experience.

We elected to communicate the existence and proximity of nearby persons in reality by inserting user silhouettes captured by a Microsoft Kinect into the VR scene in real-time. These cut-outs were acquired using Kinect tracking and depth-maps of proximate persons, and were positioned in VR at the same distance and position from the user as in reality. This direct mapping between reality and virtuality was maintained, regardless of the user's position in virtuality. The Kinect was physically positioned alongside the user, who was sat at the rear of the room, thus it captured the usable space of the room in front of the user. The aim was to provide users with an equivalent perception of their surroundings and personal space as they would have without wearing the HMD – a person 1 metre away in reality would be in the same place 1 metre away in VR.

For our low engagement, casual awareness state, the inserted proximate person was presented as a ghost-like transparent figure. For full engagement, the relevant section of reality (the proximate person) was presented fully opaque, as with the partial blending condition in the previous study.

Study Design

In this study, we aimed to establish firstly that awareness of presence and proximity could be improved through using room-based sensors, and secondly that we could communicate a reduced level of awareness for the benefit of further preserving the sense of presence, defining a low engagement state that users could persistently remain in. 12 participants were recruited from University mailing lists (Age mean=27.8, SD=4.76, 11 male 1 female).

We again hypothesised that, as we increased the amount of real feedback of proximate persons into virtuality, the user's sense of presence would diminish as s/he would suffer increased distraction from their VR experience. However, their awareness and sense of comfort with respect to personal space would increase. Our three conditions were *(1) Baseline:* The normal VR environment; *(2) Low engagement:* With a transparent, ghost-like presence in virtuality, and *(3) Full engagement:* With an opaque presence in virtuality.

Our engagement states are depicted in Figure 11. For the VR scene, the Tuscan villa was used once more. Users were given 3 minutes per condition to explore the villa whilst wearing the HMD and headphones, conditions were counterbalanced. In order to provide meaningful feedback of activity in reality, every 30 seconds the experimenter would enter the view of the Kinect and perform one of three activities for 15 seconds: writing on a whiteboard, standing using a phone, and sitting down. Sense of presence was again measured by the IPQ. For measuring the participants' awareness of social presence in reality, the "Social Presence – Passive Interpersonal" factor from the Temple Presence Inventory [12] was used. Awareness, comfort and distraction were measured using a 7-point Likert scale.

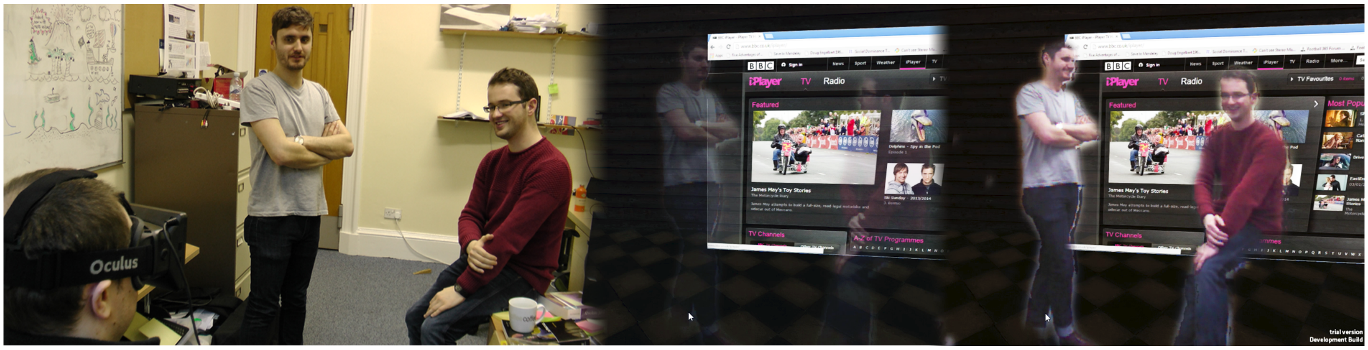

Figure 11. Our evaluated low and high engagement states. When a person enters the same physical space as the VR user, they are faded into the virtual view. When the user wishes to engage with them, they would become fully opaque. Left: reality; Middle: Low engagement; Right: High engagement.

Factor	1: Base-line	2: Low engage	3: Full engage	Friedman Test	Wilcoxon *Post-hoc*
Spatial	4.23 (0.90)	4.48 (0.72)	4.40 (0.74)	$p = 0.50$	NA
Involvement	3.92 (1.30)	3.12 (1.16)	3.23 (1.21)	$p = 0.09$	NA
Realism	3.25 (0.81)	3.46 (1.16)	3.00 (1.26)	$p = 0.14$	NA
Sense of being there	4.50 (0.91)	4.42 (0.79)	4.25 (1.29)	$p = 0.87$	NA
TPI Passive	0.92 (1.12)	2.53 (1.67)	4.56 (0.91)	$\chi^2(2) = 20.7,$ $p < 0.01$	1-3, 2-3
Distraction	1.17 (0.84)	2.92 (1.62)	3.08 (1.83)	$\chi^2(2) = 9.41,$ $p < 0.01$	1-2, 1-3
Awareness	1.42 (1.38)	3.75 (1.36)	5.42 (0.67)	$\chi^2(2) = 21,$ $p < 0.01$	1-2, 1-3, 3-2
Comfort	4.58 (1.31)	4.00 (1.65)	4.00 (1.76)	$p = 0.56$	NA

Table 3. Means (Std. deviation). IPQ, higher is better. TPI Social Presence (Passive Interpersonal), higher is more awareness. Distraction from VR due to reality, higher is more distracting. Awareness of reality, higher is more aware. Comfort in personal space, higher is more comfortable. Friedman test was conducted with *post hoc* Bonferroni corrected Wilcoxon tests.

Results

A Friedman test with *post hoc* Wilcoxon's tests was performed where applicable. Results can be seen in Table 3. There was no significant difference in sense of presence as we varied the amount of reality present. The full view of proximate persons statistically increased both distraction due to, and awareness of, reality, with the TPI Passive Social Presence questionnaire demonstrating much higher awareness of who was there in Condition 3, versus Conditions 1 / 2. However, for this scale there was no significant difference between Conditions 1 and 2. No effect was observed for comfort.

Interviews

A *post hoc* interview was held with each participant. The format was semi-structured, asking their opinions on being able to perceive reality in VR. These interviews were transcribed and coded, with the most common themes listed in Table 4.

Item Discussed	% of Participants
Wanted presence to be communicated differently	58% (7 people)
Wanted to be able to selectively bring reality in and out of view based on engagement	58% (7 people)
Thought presence of proximate persons from reality in VR scene was avatar-like or fit with virtuality	42% (5 people)

Table 4. Frequently mentioned items in participant interviews.

Participants appreciated knowing who was in the room, but frequently argued that they wanted a warning regarding the appearance of proximate persons, or that their existence be communicated more discreetly or abstractly:

> "I don't think I need to see the person; if it just popped up that person's name, like on XBox live or PS3 "this person wants to play with you"."

> "I think an example would be I'm sitting facing a monitor and I can't always tell when you come in... If they open the door (and) something came up on my monitor... I could go 'oh right' instead of having to wait for them to tap me on the shoulder"

> "I think that I'd need some cue before someone entered my view"

> "Other virtual things... butterflies, you know, other virtual things could pop up to represent somebody."

Frequently, participants expressed a wish to have control over how much reality was brought into virtuality, often selectively and based on their engagement with it:

> "I'd want to know if people are looking at me, or if they are looking at somebody else.. so if I'm told they've entered the room I could use a slider.. gradually fade it up and see who is there, what they are doing."

> "I found the opaque view quite disruptive... If there was some way where you could just do... "boom" (go away)... but then the downside would be what if they are in the room and want to talk to you in a few minutes, how do they get your attention without hitting you?"

> "There was one or two moments where I was like go away from here!"

> "If you were trying to do a task in it (virtuality) it might be annoying, so it'd be good to be able to just turn it on (or off)."

Discussion

Incorporating reality significantly improved awareness of proximate persons, at the expense of increased distraction, with our full engagement state providing vastly improved awareness (see Table 3) while allowing users to remain present in VR. However, our low engagement state, while providing a decreased awareness, failed to be significantly less distracting than our full engagement state, and thus this does not represent an ideal presentation (which we would expect would have some awareness, with significantly less distraction). Participant interviews suggested that the lesser, or more casual, state of awareness should likely incorporate non-visual representations of proximity and existence, not necessarily tied to gaze. A frequently mentioned concept was the use of textual alerts in messaging and gaming, where users are informed of the online presence of contacts. This approach was cited as being unobtrusive, while conveying the most important information: who is there and what they are doing. While our work has used the visual display, using other feedback modalities is a logical next application of our engagement-dependent augmented virtuality framework.

Interestingly, in varying the amount of reality when rendering proximate persons, markedly different effects were seen compared to interactive objects and peripherals, with no significant effects in terms of any of the IPQ sense of presence factors. Based on qualitative feedback, we suggest that whereas the existence of objects might be disconcerting or unnatural, the existence of people in a Tuscan villa was not problematic and was, in fact, in line with user expectations. This reflects Slater's conception of presence as *plausibility illusion* – if the view of reality is logically consistent with the context of the VR scene, this may preserve (or even reinforce) the sense of presence. So we could modify the appearance of reality to appear aesthetically in keeping with VR e.g. rendering an overlay on the keyboard to make it appear part of a cockpit.

While enabling engagement-dependent awareness will require future work, primarily in designing a system that can infer the engagement between the user and others, these results show how such a system might provide awareness based on engagement. We envisage an outside user attendings to the VR user, with a system inferring this engagement and incorporating this outside person into the VR scene.

CONCLUSIONS

We have shown a user-driven and experimentally validated need for the incorporation of aspects of reality into VR, specifically the ability to interact with objects and peripherals, and maintain awareness regarding proximity and existence of other people. Currently, a VR user's capability to interact with reality is significantly impaired and this directly impacts the usability of HMDs in everyday life. We have demonstrated this impact both in a survey of 108 VR HMD users, and in a controlled typing study.

Furthermore, we have presented a solution to this problem within the context of interaction with objects and peripherals: selective engagement-dependent Augmented Virtuality. Our solution enables interaction with reality, while preserving presence and immersion in VR, by selectively blending relevant parts of reality with virtuality. We do this in an engagement-dependent manner, inferring when and how much to blend reality into virtuality based on the user's engagement with objects in reality. This approach is distinct from past work, seeking to preserve presence and place illusion with a mechanism that minimises the blending of reality.

We offer selective engagement-dependent AV as a first step toward intelligently incorporating aspects of reality into the VR experience. Through this work, VR HMD users will be able to perform necessary interactions, such as taking a drink, using a peripheral, or being aware of whether they are alone or not, all without having to leave their VR experience, or remove their HMD.

ACKNOWLEDGMENTS

This work was supported in part by Bang & Olufsen, the EPSRC and the Danish Council for Strategic Research of the Danish Agency for Science Technology and Innovation under the CoSound project, case number 11-115328. This publication only reflects the authors' views.

REFERENCES

1. Barrett, J., and Krueger, H. Performance effects of reduced proprioceptive feedback on touch typists and casual users in a typing task. *Behaviour & Information Technology 13*, 6 (1994), 373–381.
2. Benford, S., et al. Understanding and constructing shared spaces with mixed-reality boundaries. *ACM TOCHI 5*, 3 (1998), 185–223.
3. Billinghurst, M., and Kato, H. Collaborative mixed reality. In *Proc. ISMR '99*, ACM (1999), 261–284.
4. Bruder, G., and et al. Augmented virtual studio for architectural exploration, 2010.
5. Carvalho, F. G., et al. Toward the design of transitional interfaces: an exploratory study on a semi-immersive hybrid user interface. *Virtual Reality 16*, 4 (2012).
6. Davis, L., et al. Enabling a continuum of virtual environment experiences. *IEEE Comput. Graph. Appl. 23*, 2 (Mar. 2003), 10–12.
7. Gonzalez, G., et al. Evaluation of text input techniques in immersive virtual environments. In *New Trends on HumanComputer Interaction*, J. A. Macas et al., Eds. Springer London, 2009, 109–118.
8. Grasset, R., Lamb, P., and Billinghurst, M. Evaluation of mixed-space collaboration. In *Proc. ISMAR '05*, IEEE (2005), 90–99.
9. Hutchison, A. Back to the holodeck: New life for virtual reality? In *Proc. DIMEA '07*, ACM (2007), 98–104.
10. Kaber, D. B., and Zhang, T. Human factors in virtual reality system design for mobility and haptic task performance. *Reviews of Human Factors and Ergonomics 7*, 1 (2011), 323–366.
11. Kristensson, P. O., and Vertanen, K. Performance comparisons of phrase sets and presentation styles for text entry evaluations. In *Proc. IUI '12*, ACM (2012).
12. Lombard, M., et al. Measuring presence: the temple presence inventory. In *Proc. ISPR '09* (2009).
13. Metzger, P. Adding reality to the virtual. In *Proc. VR '93*, IEEE (1993), 7–13.
14. Milgram, P., and Colquhoun, H. A taxonomy of real and virtual world display integration. *Mixed reality: Merging real and virtual worlds* (1999), 5–30.
15. Milgram, P., and Kishino, F. A Taxonomy Of Mixed Reality Visual Displays. *IEICE Transactions on Information and Systems 77*, 12 (1994), 1321–1329.
16. Pohl, H., and Murray-Smith, R. Focused and casual interactions: Allowing users to vary their level of engagement. In *Proc. CHI 2013*, ACM (2013).
17. Rekimoto, J., and Nagao, K. The world through the computer: Computer augmented interaction with real world environments. In *Proc. UIST '95*, ACM (1995).
18. Schubert, T., Friedmann, F., and Regenbrecht, H. The experience of presence: Factor analytic insights. *Presence: Teleoperators and virtual environments 10*, 3 (2001), 266–281.
19. Slater, M. Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Phil. Trans. R. Soc. Lond. B: Biological Sciences 364*, 1535 (2009), 3549–3557.
20. Soukoreff, R. W., and MacKenzie, I. S. Metrics for text entry research. In *Proc. CHI '03*, ACM (2003), 113–120.
21. Steinicke, F., et al. A virtual body for augmented virtuality by chroma-keying of egocentric videos. In *3DUI 2009*, IEEE (2009), 125–126.
22. Tecchia, F., et al. I'M in VR!: Using Your Own Hands in a Fully Immersive MR System. In *Proc. VRST '14*, ACM (2014), 73–76.
23. Thomas, B. H. A survey of visual, mixed, and augmented reality gaming. *Computers in Entertainment 10*, 3 (2012), 3:1–3:33.
24. Van Baren, J., and IJsselsteijn, W. Measuring presence: A guide to current measurement approaches. *OmniPres project* (2004).