# Lost in the Rift: Engaging with mixed reality

**Virtual reality users are torn between the real and virtual worlds. Determining how, and when, to show elements of reality in a virtual view is key to providing usable VR experiences.**

*By Daniel Boland and Mark McGill*

DOI: 10.1145/2810046

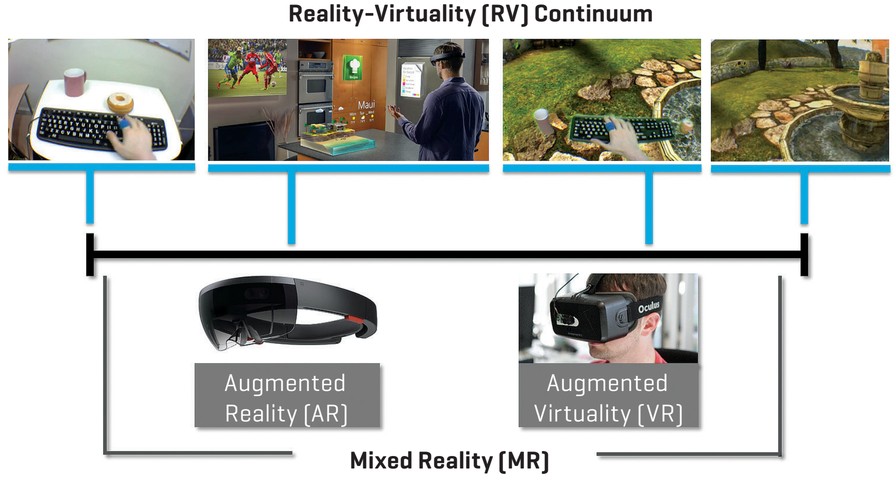
A new virtual reality (VR) user reaches out and feels around in the dark, searching for the controller yet unable to see where it is. Elsewhere, immersed in their own VR experience, another user is oblivious to the other person who walks into the room; they are standing side by side, but unaware of the other's presence. Mixed reality approaches that allow these real-world elements to be blended into VR can break the user's sense of immersion in their virtual experience. Our research on engagement-dependent mixed reality has solved this problem, by selectively blending real elements into the virtual world as users wish to engage with them, which creates a seamless interaction across the continuum of real and virtual reality.

The Oculus Rift's Kickstarter campaign triggered a resurgence of interest in VR head-mounted displays (HMDs). Advances in small form factor displays (e.g., the high refresh rate, low persistence, and high definition panels typically used in mobile devices) demonstrated high fidelity VR HMDs were now not only technologically feasible, but a viable and affordable consumer reality. What followed has seen the likes of Samsung (Gear VR), Sony (Morpheus), HTC/Valve (Vive), Oculus/Facebook (DK1/2, CV1), and Google (Cardboard) battling to be the leader in this VR renaissance. But there remains a number of sizeable problems in trying to deliver a VR experience that is usable in the real "consumer" world. Some of these problems are fundamentally technological. For example, the fidelity of the VR experiences increases—as our capability to render and display them does—through more powerful GPUs, better displays, and wider field-of-view lenses. Simulator sickness, another major problem for users, is being addressed with additional sensing such as external tracking cameras combined with headset-based inertial motion sensing, allowing for every movement in reality to be mapped to VR. However, perhaps the most pressing concern currently facing the VR community is not related to the rendering or display of VR experiences, but instead how we interact with VR experiences and make VR HMD usage compatible with real-world environments.

## THE GLOVES COME OFF

The problem immediately becomes apparent when you first wear a VR HMD. You find yourself in a virtual world with no perception of reality. How do you interact with this new virtual world you inhabit? The go-to in-

**Figure 1: Milgram et al.'s reality-virtuality continuum. Here we can see augmented reality (pictured the Microsoft Hololens) and augmented virtuality (pictured the Oculus Rift DK2) contained within the scope of mixed reality, with examples given of what the user might perceive at different points along the continuum.**




put mechanisms (game controllers, mice, and keyboards) immediately become more difficult to use when they can't be seen. If you want proof of this, close your eyes, stand up, spin around a bit on the spot, and then try to type something on your keyboard. If you survive unscathed, you might remark that it is anything but easy. There are styles of interaction that don't need peripherals, such as gestural interaction. But even this can be hazardous when you can't see the real obstacles that surround you in your household, such as plant pots, walls, or beloved family pets. One way of solving this problem is to develop highly targeted interactions for specific VR contexts, like the flight rig in the Birdly VR Simulator.[1] General consumer use of VR will, however, require generic and adaptable interaction techniques, suitable for use in home and office spaces without significant deployment or installation.

As such, considerable efforts have gone into recreating real-world interactions in VR. Hand tracking (e.g., using the leap motion VR) allows users to interact with 3-D VR elements in a natural manner, using their proprioception. Being able to reach out in reality and touch objects virtually allows common interactions without the go-between of a peripheral. However the word "touch" is problematic here. Hand-tracking solutions often lack haptic feedback when touching a VR object, which pulls users out of the virtual experience, making precise manipulation difficult.

This lack of haptic feedback can be compensated for in numerous ways. For example, Neurodigital's Gloveone[2] uses vibrotactile actuators embedded in a glove, allowing a limited perception of virtual objects. In contrast, the Dexta Robotics Dexmo[3] uses mechanical actuators to create a force-feedback exoskeleton glove, allowing virtual objects to be grabbed and released. Mid-air feedback also shows promise, using ultrasound [1] or air vortices [2]. Given the relative lack of maturity of these technologies, interaction with real tangibles in some form is likely to be required for a variety of experiences.

**In shutting off reality, we can become extremely immersed in VR without distraction. But the more we shut off reality, the more difficult certain interactions become.**

With respect to tangible objects, handheld motion controllers (e.g., Oculus Touch, Playstation Move, and SteamVR) are the most promising current approach. They typically provide the capability for high-bandwidth input and "hand presence," meaning the positional tracking of hands and/or fingers, at the cost of being able to naturally manipulate virtual objects through touch. Other approaches aim to retain the capability to pick up and manipulate virtual objects. Substitutional reality appropriates real-world tangible objects to represent virtual objects, for example turning a torch into the hilt of a VR lightsaber [3]. Apart from these, other modalities for interaction are also available. Voice, gaze, and motion tracking all allow some semblance of mapping the actions of the real world to the virtual. Allowing the user to move in VR can be achieved using omnidirectional treadmills, external sensing with redirected walking, substitutional reality on a larger scale, or even room-scale VR [4].

However, many of the discussed interaction techniques share a common problem; the interaction starts in reality, yet the VR user has limited perception of that reality.

## LOST IN THE RIFT

Being unable to perceive reality when in VR has serious consequences regarding usability. Take handheld controllers; if we put the controller down, we need to be able to see where it is in order to pick it back up again. For hand-tracking interac-

1 http://somniacs.co/

2 http://www.gloveonevr.com/

3 http://www.dextarobotics.com/products/Dexmo

tions, we need to be aware of what real-world objects might be interacted with. Waving one's hands in the air may make for an immersive experience, but knocking over a cup of water would make for a rather moist one. For movement through real-world spaces, we need a perception of where we can move. Walking into a wall impacts your immersion in VR as well as your head. The problems don't stop there. There are a variety of scenarios where a user might want to perform a necessary interaction with reality (e.g., typing on a keyboard or eating some popcorn), or might need to be aware of reality (e.g., your computer spontaneously combusting). Not only can this make for awkward situations, it can also leave users feeling vulnerable with little control over their personal space.

In a survey of 108 existing VR HMD users, we confirmed there were significant problems regarding interacting with peripherals and communicating the social context of the room. Users reported being ineffective at interacting with real-world objects and peripherals. This was unsurprising, but confirmatory, given that users were unable to see these objects and peripherals. For social contexts, we found users were unaware of the presence and proximity of others in the room. In both cases, users agreed the VR HMD should have in-built support for incorporating reality.

In essence, the isolation a VR HMD provides is both its greatest strength and most significant weakness. In shutting off reality, we can become extremely immersed in virtual reality without distraction. But the more we shut off reality, the more difficult certain interactions become. The problem space is in deciding when, and how, to breach that isolation and communicate information regarding real-world context and objects.

The extent to which we breach this isolation is defined by where we are on the "reality-virtuality (RV) continuum" [5], also known as the "mixed reality continuum" (see Figure 1). The RV continuum is a scale ranging between completely real (meaning the user can see nothing but reality) and completely virtual (meaning the user can see nothing but VR), with all the points in-between comprising of mixed reality—a blend of real and virtual. Depending on how this mixing is skewed, it might be interpreted as augmented reality (AR), a view of reality that is augmented with virtual aspects, or augmented virtuality (AV), a view of VR that is augmented with aspects of reality. Early work in this area by Metzger [6] proposed a seamless integration of real-world human interfaces with the virtual world, and our work explores how to realize this.

**There will be HMDs in the future that will support both AR and VR/augmented virtuality modes and will have the ability to transition between any point on the RV continuum.**

In order to augment virtuality with reality and achieve this seamless integration of real and virtual, we first need a VR HMD that is able to sense elements of reality. Such sensing can be room-wide (e.g., Oculus Constellation, HTC/Valve Lighthouse, or Microsoft Kinect) or mounted on the headset itself (e.g., Leap Motion VR, Gear VR camera, or Google Project Tango). These sensing techniques typically support depth capture, person identification, object/hand tracking, and biometrics across a wide field-of-view. Using multiple sensors, we can start to accurately map out the room. For example, we can track the VR user (e.g., position and orientation, compensating for drifts in on-board inertial sensing), the social context of the room (e.g., who is there, what are they attending to, and where they are), and the physical context of the room (e.g., objects, obstructions, and walls). Our work was predicated on these sensing capabilities being widely available, and was concerned with how we might use this sensing to enable the VR HMD user to interact with, and be aware of, reality. To do this, we had to answer two questions: How should we incorporate reality and when?

**Figure 2. Feedback from real objects can be used to augment virtuality, providing a low-level control loop that can support high bandwidth interaction, such as typing. The more the user engages with reality, the more real feedback is mixed with virtuality.**

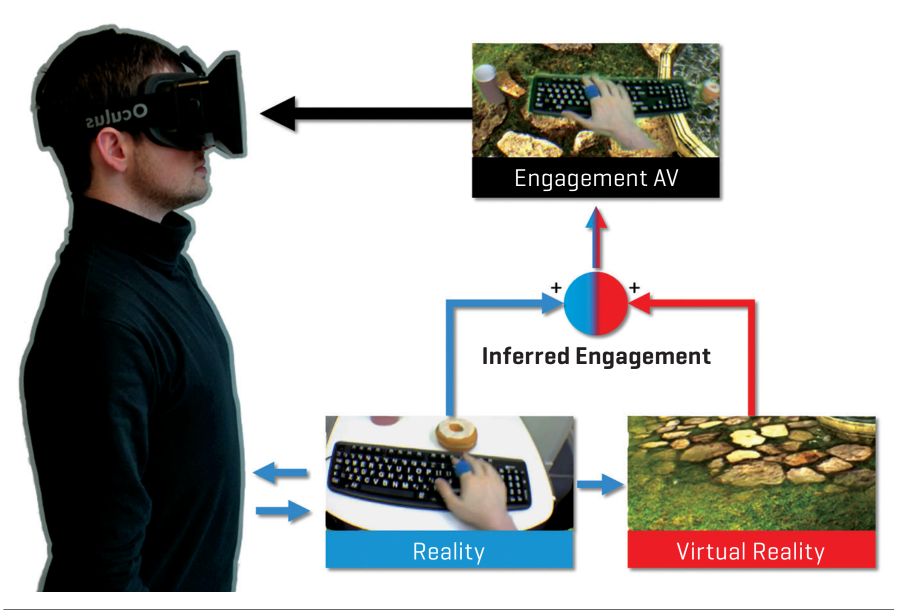


**Figure 3. A VR HMD wearer's view of engagement-dependent mixed reality. Here, the user has reached out to interact with an object; in turn, the system has inferred this engagement and temporarily brought the available objects to be interacted with into the VR view.**

### ENGAGEMENT-DEPENDENT MIXED REALITY

There are clearly scenarios where users would want to blend real elements into their virtual experience. The question then is how to find the appropriate balance between real and virtual—where does the user want to be on the RV continuum? Our work explores how to intelligently infer which elements of reality or virtuality the user is engaging with, and adapts the mixed-reality blending accordingly (for more details beyond this article, see McGill et. al. [7] and the associated video[4]). This engagement is considered to be the user's attention to an element, such as by a sustained gaze, or the user's control of an object reaching for a cup or typing on a keyboard. This "engagement-dependent" mixed reality allows a number of pressing VR usability issues to be addressed.

Let us again look at the case of the keyboard or controller in detail. We know finding and interacting with these peripherals unsighted is difficult. We could incorporate these objects into mixed reality; however having such real elements permanently visible in VR would impact immersion—even the most immersive scenes would be spoiled by a ghostly, hovering keyboard. By only including an object as the user engages with it (for example by reaching out for it, or by gaze), the engagement-dependent mixed-reality approach enables a high bandwidth interaction such as keyboard input while allowing users to otherwise remain immersed in their VR experience (see Figure 2).

As we infer the user's engagement, we can also blend elements where there is a high likelihood of engagement. Obvious examples include necessary objects, such as a cup that is frequently drank from (see Figure 3), or hazards such as a wall in the path of the user (as is the case with the feedback provided by the HTC Vive headset).

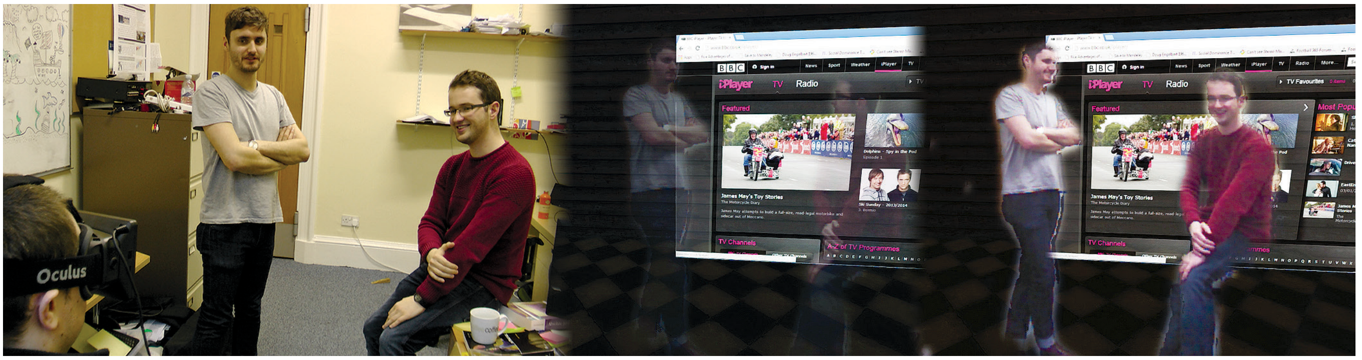
This could also apply in virtuality, for example incorporating an avatar of a friend into an otherwise real-world view. This is particularly useful for blending people into the VR view as they enter the room. It can be assumed a VR user wants to be aware of people entering their personal space, and so there is an implicit level of engagement with these people always blended in slightly as "ghosts." As the VR user interacts with others, they can be blended in more fully, occluding part of virtuality (see Figure 4). This example, with its degrees of blending, demonstrates how engagement-dependent mixed reality spans the RV continuum dynamically.

### THE CONVERGENCE OF VR AND AR HMDS

While much of our work is based around the concept of a VR HMD combined with sensors for capturing

4 www.youtube.com/watch?v=SHdfxuh7_GY

**Figure 4. 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.**



and incorporating reality, there is another class of mixed-reality displays currently in use. AR headsets typically rely on a see-through display where virtual content can be rendered on top of the user's view of reality (as opposed to reality rendered virtually). They occupy the AR part of the RV continuum, with Microsoft Hololens being a recent example of the state of the art.

Because AR headsets have the ability to transition between a range of points on the RV continuum, we can apply our engagement-dependent mixed-reality approach. This time the blending would govern when and how much virtual content is rendered on top of reality. For example, a lingering gaze on a tourism landmark might be a cue to incorporate additional relevant information.

The underlying technologies used in AR headsets suggest in the future there will be HMDs that support both augmented reality and VR/augmented virtuality modes, and will have the ability to transition between any point on the RV continuum. For example, we can envisage Hololens–like headsets where there is an additional display layer that can selectively occlude reality, such that the rendering of virtual content can be done on top of reality or on top of a blank, dark canvas—much like how 3-D TV active-shutter glasses work today. Such headsets would be at once empowering and isolating in equal measure, supporting instantaneous transitions between virtual spaces and augmented reality spaces. But how would we make these transitions seamless? What elements of reality should get carried back and forth? And how do we provide users with a stable mental model for interacting with reality as their immersion into VR increases? There would be a need for consistent behaviors and rule sets regarding transitions across the RV continuum, be they pertaining to interactions with reality, awareness of social contexts, and so on. We suggest an engagement-dependent mixed-reality model could underpin such interactions, and such a model is a necessity if interactions with reality are to be as seamless and effortless as users might expect.

## CONCLUSIONS: CAN I SEE YOU NOW?

What we've described is a means to balance the needs of VR users who are torn between two worlds: reality and virtuality. Users want to interact with a virtual world, yet have to route said interaction via the real world. This balance between immersion in virtuality and awareness of reality is key to providing usable, yet immersive, VR headsets that can be used in homes, and offices, as well as other shared, social, complex spaces without impediment. Our work provides a foundation for both AR and VR displays, with transitions in mixed reality governed by the user's engagement. This engagement-dependent approach to mixed reality provides experiences that better adapt to user needs, and break down the isolating digital and physical barriers raised by head-mounted displays.


**References**

[1] Wilson, G., Carter, C., Subramanian, S., and Brewster, S. Perception of ultrasonic haptic feedback on the hand: localisation and apparent motion. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems* (CHI '14). ACM, New York, 2014,1133-1142.

[2] Sodhi, R., Poupyrev, I., Glisson, M. and Israr., A. AIREAL: interactive tactile experiences in free air. *ACM Trans. Graph.* 32, 4 (July 2013).

[3] Simeone,A., Velloso, E., and Gellersen, H.. Substitutional Reality: Using the Physical Environment to Design Virtual Reality Experiences. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15). ACM, New York, 2015, 3307-3316.

[4] Machkovech, S. SteamVR: The room-scale VR world that feels like an "IMAX in your house." June 13, 2015. *Ars Technica UK*. http://arstechnica.co.uk/?p=19673.

[5] Milgram, P., and Fumio, K. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems* 77.12 (1994), 1321-1329.

[6] Metzger, P.J. Adding reality to the virtual. In *Proceedings of the 1993 IEEE Virtual Reality Annual International Symposium* (VRAIS '93). IEEE Computer Society, Washington, DC, USA, 1993, 7-13.

[7] McGill, M., Boland, D., Murray-Smith, R., and Brewster, S. A Dose of Reality: Overcoming Usability Challenges in VR Head-Mounted Displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15). ACM, New York, 2015, 2143-2152.


**Biographies**


Daniel Boland applies data science and machine learning approaches to HCI. This work was informed by his Ph.D. thesis on adapting to user engagement. Learn more at www.dannyboland.com.

Mark McGill is a Ph.D. student at the University of Glasgow in the Multimodal Interaction Group, led by Prof. Stephen Brewster. Outside of VR, his research interests are in multi-user and multi-view TV. Learn more at www.markmcgill.co.uk.





1528-4972/15/09 $15.00