Science Arts & Métiers (SAM)

is an open access repository that collects the work of Arts et Métiers Institute of Technology researchers and makes it freely available over the web where possible.

To cite this version :

William Eric MARSH, Frédéric MERIENNE - Nested immersion: Describing and classifying augmented virtual reality - In: IEEE Symposium on 3D User Interfaces- -Workshop, France, 2015-03-23 - IEEE Virtual Reality Workshop on Perceptual and Cognitive Issues in Augmented Reality (PERCAR) - 2015

Nested Immersion: Describing and Classifying Augmented Virtual Reality

William E. Marsh[*]   Frédéric Mérienne[†]

Arts et Métiers ParisTech — CNRS Le2i, Institut Image

ABSTRACT

We present a system, intended for automotive design review use cases, that incorporates a tracked tablet in a CAVE, where both the tablet and the CAVE provide different views and interaction possibilities within the same virtual scene. At its core, this idea is not novel. However, the literature reveals few examples of this paradigm in which virtual information is presented on a second physical device to augment an immersive virtual environment. Similarly, it is unclear where the system should be positioned within existing augmented/mixed/virtual reality taxonomies. We argue that interactions occur within a nesting of virtual and physical contexts, and that formalizing these relationships is important when attempting to understand perceptual issues. The goal of this paper is, thus, to describe the new system by proposing a scheme to formally identify sources of bias and then adapting an existing taxonomy to classify such systems.

Index Terms: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities; H.5.2 [Information Interfaces and Presentation]: User Interfaces—Input devices and strategies; H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces — Computer-supported cooperative work

1 INTRODUCTION

Virtual reality (VR) is becoming increasingly common for design review applications in the automotive industry. Because these systems are used by diverse groups of stakeholders with a wide range of technology skill sets, intuitive interfaces are needed for manipulating and modifying the virtual models. Many of these users are familiar with interfaces found on a common smartphone or tablet and an augmented reality metaphor makes sense as a way to peer into a virtual scene. For this reason, a system was developed that allows users to take a tracked tablet into a CAVE and interact with the same immersive virtual scene on both display devices.

Such systems are immersive, designed to evoke a high sense of virtual presence, the subjective sense of "being" in a virtual environment (VE) [16]. Awareness of physical system attributes can lessen this experience, but this physical awareness is paradoxically necessary for effective interaction with a tablet display. This may have perceptual implications, affecting not only system usability but also perception in the VE. Several taxonomies exist to categorize mixed reality and, more generally, multi-display systems. However, none adequately acknowledge the nesting of immersive environments and relationships between the physical and virtual contexts in such a system. These relationships relate to presence and perception, so filling this gap in the literature is an important first step in formally identifying and comparing sources of perceptual bias in such systems.

---

[*]e-mail: william@wemarsh.com
[†]e-mail: frederic.merienne@ensam.eu

Figure 1: VARI[3] allows a user to explore and modify a single virtual vehicle model displayed simultaneously on both a CAVE and a tracked tablet.

1.1 VARI[3]

A CAVE-based design review system known as VARI[3] (Virtual & Augmented Reality Interactive and Intuitive Interface), initially intended for applications in the automotive industry, has been developed. As shown in Figure 1, the CAVE displays a virtual car model, which is augmented with additional information on a tracked tablet, an iPad Mini. Though VARI[3] may be used in multiple physical settings, the CAVE of interest in this work is four-sided, measuring 3×3 m horizontally and 2.67 m in height, with passive stereoscopic rear projection on three walls and front (top) projection on the floor. The iPad Mini has a 20 cm (diagonal) touchscreen display. It and the user's head are continually tracked with an ARTTRACK2 optical tracking system, with the head position used to update the center of projection in each frame.

While viewing the virtual model in the CAVE, the user holds the tablet which provides an additional view into the scene and facilitates interactions. The user can use the tablet to perform tasks while moving freely within the CAVE boundaries. The following is a non-exhaustive list of VARI[3] functionality.

Displacement: As virtual models are typically displayed at a 1:1 scale, and because a typical car is larger than the physical CAVE area, the physical and virtual environments cannot be fixed relative to one another. The tablet interface allows the user to translate and rotate the CAVE in virtual space (or, stated differently, translate and rotate the VE around the CAVE). In this way, a user can inspect the model from particular angles of interest.

Annotation: To flag problems or add other comments to a virtual model, the user can use the tablet to directly select the location to be annotated and enter text using a virtual keyboard.

Changing components: Virtual components can be swapped using the tablet interface. For example, in an automotive design

application, multiple dashboards can be compared to check exterior visibility differences.

Cut plane positioning: The user can position and orient a cut plane using the tablet interface, allowing for a cutaway view of the virtual model.

Virtual mirror: A mirror mode allows the user to view the back of a virtual object, such as under the dashboard of a virtual car. Additionally, this allows the tablet to function as a rear-view mirror, as in a real car.

Note that VARI³ allows extensive read/write access from the tablet into the same VE that is viewed on the immersive CAVE displays. All of the above functions require accurate spatial perception for effective use, so it is important to understand the sources of bias in the system.

1.2 Existing taxonomies

Some past literature has attempted to classify systems that allow for viewing related information on more than one display at a time, though none consider the interplay between physical and virtual aspects of immersive VR systems. For example, Terrenghi, Quigley, and Dix [15] developed a taxonomy of *coupled* display devices. In their work, they are concerned with defining what it means for displays to be coupled, the fluidity of these couplings, and the types of social activities that are mediated by the system. Most relevant to the present discussion is the definition of coupling, which requires an interaction to have shared input and/or output, and related states to be linked at the system level (not only in the mind of the user).

Focused on the costs of attention-switching between displays, Rashid, Nacenta, and Quigley [10] created a taxonomy of multi-display user interfaces. They identified five relevant factors: display contiguity, angular coverage, content coordination, input directness, and input–display correspondence. They did not, however, consider immersive VR systems in which a user experiences presence in a VE that is split or replicated between physical displays.

Milgram, Takemura, Utsumi, and Kishino [7] proposed the frequently cited reality–virtuality (RV) continuum, along which one can classify environments as completely real, completely virtual, or somewhere in between. Such an environment that falls in between the extremes is termed *mixed reality* (MR), defined as "one in which real and virtual objects are presented together within a single display." Within MR, the following classifications were defined, both of which are relevant to the present discussion.

Augmented Reality (AR): A real environment is augmented by virtual artifacts and information.

Augmented Virtuality (AV): A VE is augmented by real artifacts and information.

In practice, a monitor-based or see-through display in an AR system can generally only present a *read-only* view into the *real world* and, if modification (*writing*) were allowed, it would only be possible on a clearly delineated subset of scene objects and attributes, selected in advance. Note that the RV continuum does not explicitly allow for a VE to be augmented with virtual information, and it is unclear where on the scale such a system would lie.

Stoev and Schmalstieg [13] conceived of a taxonomy of *through-the-lens* techniques, in which a distinction is drawn between two concurrent viewpoints. In this framing, the user is situated within what we, for consistency in this paper, will call the *primary environment* and sees it from the *primary viewpoint*. While in this primary environment, the user can then look through an *output window* to view what we will call the *secondary environment* from the *secondary viewpoint* (in virtual space, the output window is called the

viewing window). Importantly, the primary and secondary environments are both considered to be virtual, with one being defined as a copy of the other, and both are assumed to be rendered on the same display device. The taxonomy allows for the states of the output window and the secondary environment to be classified. The following states are possible for the output window. For more information, see Figure 2 in the original paper [13].

Case O1: At a fixed position and orientation in the primary environment.

Case O2: At a fixed position in the image plane and oriented so the user can always see it.

Case O3: Mapped onto the location of a handheld pad, which can be freely positioned and oriented by the user.

In addition, the following possible states are defined for the secondary environment. For more information, see Figure 3 in the original paper [13].

Case V1: Fixed relative to the primary environment.

Case V2: Fixed relative to the viewing window.

Case V3: Fixed relative to the primary viewpoint.

1.3 Augmented reality

Augmented reality is a large, diverse, and rapidly growing research area. This review is limited to just two examples of past systems that can be considered analogous, in some important respects, to the VARI³ system described above in Section 1.1.

Some past researchers have employed a lens metaphor to "look into" the real world. Looser, Billinghurst, and Cockburn [5] developed such a system, which allows users to augment a physical book using a semi-transparent head-mounted display which acts as a "magnifying glass." The book is readable as a normal book, but the lenses add supplemental virtual content. Virtual scenes pop out of the pages or users can instead choose to explore the content from a first-person perspective.

Szalavári and Gervautz [14] developed the handheld *Personal Interaction Panel* (PIP), comprising a tablet and a pen prop. Both the panel and the pen were "dumb," incorporating technology for realtime position and orientation tracking but with no computational or display capabilities of their own. The PIP was augmented along with the real world, using the same display device, a see-through head-mounted display. The output window in the PIP system falls under Stoev and Schmalstieg's [13] case O3 because the "panel" can be thought of as a movable "pad" in the primary world. The state of the secondary environment is not constrained by hardware choices and thus it can vary depending on the needs of each individual use case.

2 Computer-Mediated Virtual Reality

Few systems exist that use mobile devices to augment VR, and those that do generally do not attempt to capitalize on the benefits of immersion. Some systems incorporate a tablet allowing a user to magnify or extend on information seen on a large display. Using these systems often involves holding a tablet up to a screen [11] or a tabletop [12], as opposed to moving freely throughout the tracked space of a larger immersive VR system. This usage model implies a low level of virtual presence, because a user must perform actions relative to physical hardware.

In a notable example of a system that does benefit from immersion, Aspin [1] describe a CAVE-based system in which an immersive VE is augmented with a tracked tablet display. The tablet provides a more detailed view into the VE displayed on the CAVE walls, while still allowing natural movement through CAVE-space.

This use of multiple heterogeneous displays in immersive VR spurs unique perceptual questions. Because virtual presence is a goal, these questions may be somewhat different than would normally be encountered in MR systems.

2.1 Multiple displays

Cauchard et al. [2] considered performance issues due to visual separation in mobile multi-display environments, in which a user holds a mobile device while also interacting with one or more additional displays. As with the taxonomy of Rashid et al. [10] mentioned above in Section 1.2, this work focuses on switching costs when directing visual attention from one display device to another. They found increased performance if two displays were positioned in the same field of view. Their study was not conducted in an immersive VE, so all attention shifts occurred relative to the physical environment.

When a user is immersed in a VE while using a VR system with multiple displays, it is unclear if these attention shifts occur in virtual space or physical space. Therefore, switching costs may not be the same as those identified in the physical world.

2.2 Presence and affordances

When developing VR systems and applications, a common goal is for users to experience some degree of subjective presence in the virtual world, at the expense of subjective presence in the physical world. Normally, when using an input device, such as a gamepad, the user is still observing (with vision, but also possibly other senses) the virtual world. If the user is not looking at the gamepad then there is often no (visual) cue to explicitly break virtual presence. However, if a tablet is intended to enhance the experience by augmenting the virtual world, the user must attend to the physical device which, paradoxically, may negatively impact virtual presence.

There is a clear link to the concept of affordances and, more importantly, perceived affordances [8]. In normal VR use cases, affordances may be physical or virtual, depending on the realm in which actuation occurs. A virtual floor affords virtually stepping and a virtual handle affords virtually grasping, while the physical floor affords physically walking and a physical gamepad button affords physically pressing. Assuming the (naive, but common) goal of 100% virtual presence, all perceived affordances are virtual. Potential button-presses on a gamepad are necessarily perceived as virtual, as the physical world no longer conceptually exists when a user experiences 100% virtual presence. A related philosophical question is whether the user should perceive interaction as being with the virtual buttons, or if potential presses should be perceived as virtual actions such as "step forward."

2.3 Perceptual biases

There are many known, but often poorly understood, perceptual biases in VR. One commonly reported example is underestimation of virtual distances relative to equivalent distances in the physical world [4]. These biases might relate to impoverished graphical cues, interface attributes such as locomotion effort [9], or incidental characteristics of the physical display technology such as optical accommodation [3] or salient borders [6].

It is unclear how biased spatial models of VEs are integrated when a common scene is viewed (or interacted with) on two heterogeneous display devices. In particular, a tablet display is unlikely to be stereoscopic, so one depth cue is eliminated that would often be available on CAVE displays. Additionally, the field of view of a tablet display is very small, though it varies depending on how the user holds the tablet. This differs from a CAVE, in that the CAVE field of view is often eye-limiting regardless of a user's distance from the screens (unless the user turns around). The way that the

user holds the tablet also affects optical accommodation and vergence, both of which are also different on CAVE displays. A user must reconcile these conflicting sources of spatial information.

Much of this commentary also applies to traditional AR systems, though biases are often greater when viewing scenes in a CAVE than in the physical world. Additionally, things differ greatly when the user is not constrained to a static location. The user can move into a new physical position in CAVE-space or possibly into a virtual position using a displacement interface on the tablet. This often allows for greater possibilities than in the physical world. Systems can even conceivably allow for displacement above the ground, which may severely affect the estimation of distances. This ability to virtually displace is not generally a consideration in traditional AR systems, and it comes with challenges. A user might see a virtual location on the CAVE displays in which he would like to be standing or perhaps a viewpoint he would like to take, but he must use the tablet interface to select the destination. The location may conceivably be selected using a top-down schematic view or direct selection. In either case, the user may have trouble accurately making the selection, which requires spatial reasoning and perspective taking, together with an understanding of the relative physical positions of the display devices. Likewise, a tablet may allow for interaction with the virtual model, functionality which has no analogue in traditional AR systems. In order to use the interface effectively, a user must understand the effects of tablet-based actions on stimuli in the CAVE. For all of these reasons, both the physical and virtual contexts in which a user is immersed, as well as individual differences, may affect a user's ability to effectively interact with and move through a VE viewed on multiple heterogeneous displays.

Active manipulation of display configuration may also benefit users, improving spatial perception and usability. Seeing the same scene on multiple devices provides more flexible viewpoint and parallax control, and more information can be presented overall. But the question then is if users will have a sufficiently accurate mental model of the relationship between virtual and physical space in order to effectively use the system to realize those benefits. As above, this may depend greatly on contexts of immersion and individual differences.

2.4 Classifying VARI^3 within existing taxonomies

The VARI^3 system is unusual in that virtual information on one physical display augments the immersive VE seen on another physical display. The design is similar to that described by Aspin [1], but this system provides a much wider range of functionality. We will argue that this configuration fits the definition of neither VR nor AR, and that it in fact fits poorly into existing taxonomies.

First, one might broadly note that the VARI^3 displays are coupled, as defined by Terrenghi et al. [15]. Their work is relevant to this discussion as VARI^3 is intended for use by multiple co-located users, but here we choose to consider primarily the immersive VR aspect. Traditionally, coupled displays are considered according to their relationships in the physical world, but in a highly immersive system the user may not always be aware of the physical aspects of one or both screens. If complete immersion (100% presence) is achieved, displays should not seem coupled, because users are unaware of their very existence.

A reasonable next step is attempting to position VARI^3 along Milgram's RV continuum. In VARI^3, we are augmenting a virtual environment with virtual content. In terms of the intended immersive experience, there is no real-world component. Thus there is no doubt that VARI^3 involves VR, but this definition fails to account for the added flexibility and functionality of the tablet interface. It also fails to account for limitations due to the tablet manipulations being constrained by the laws of physics. The concurrent use of two physical displays, with different specifications and positioning, has perceptual implications that are related to, yet distinct from,

those in VR.

It may be tempting to broadly classify VARI^3 as MR, or more specifically as AR, but in terms of the functions described in Section 1.1, there is no "reality." Content on a given display is not "mixed." All displayed objects are virtual and the only links to reality are the displays themselves. In VARI^3 the user experiences presence in the virtual scene displayed on the CAVE walls. With high levels of virtual presence, awareness of the physical system is minimized.

Similarly, VARI^3 cannot be classified as an AV system, because the tablet displays only virtual content. Traditional AV systems augment a virtual scene with content from the physical world, such as the user's tracked hand. Note that a similar system could conceivably use a tablet to provide information related to the virtual and physical worlds simultaneously. For example, imagine a hypothetical displacement interface in which the tablet displays an image of the CAVE, which can be dragged to remote locations in the virtual world. Because physical objects would be used to augment a virtual world, such a system would fit the description of AV.

Finally, we attempt to fit VARI^3 into Stoev and Schmalstieg's taxonomy [13]. The output window clearly falls into case O3, because it is a mobile handheld display. As mentioned above, the VARI^3 tablet is similar to the handheld pad in the PIP system in terms of interaction possibilities, but not in terms of perception, due to the use of two distinguishable display devices. Recall that the PIP system incorporated a dumb tablet allowing for position to be tracked, but the associated graphics were displayed on a see-through head-mounted display, the same device used to display all other virtual objects in the scene. The secondary environment (that seen through the tablet) in VARI^3 can be classified into different V states depending on the current mode of operation. In normal see-through operation, the secondary environment can be classified as Case V1 because the two environments are fixed with respect to one another. It is also conceivable that systems like VARI^3 may allow for unlinking the coordinate systems for other reasons, such as a magnification or perspective-taking function. When using the displacement function, a teleport destination is chosen from a top-down schematic view. This map view is fixed regardless of the user's viewpoint or the tablet position and orientation. Viewing the tablet from an angle does not change the displayed schematic. This means that the displacement function is a special case of cases V2 and V3, in which the tablet output window displays a 2D projection of the secondary environment. Because this projection is at the same depth as the tablet display, it is not possible to view different parts of the secondary environment by viewing the tablet from an angle.

A limitation of this taxonomy is that it does not consider the physical context of viewing and interaction. Both the primary and secondary environments are virtual. We view this as a limitation in general, as the physical context always relates to perception, but it is more relevant in the case of VARI^3 because the tablet and the CAVE are both physical display devices and they are therefore related to one another in physical space. Any perceptual biases that exist in the physical environment, in addition to virtual biases, will affect perception of the virtual environment. This consideration is more important when there are two display devices, because a user's understanding of the physical relationship between the devices (orientation, position, etc.) may be biased. If we are to consider the interplay between the physical and virtual environments, then further work is necessary.

3 VISUALIZING THE CONTEXT OF IMMERSION

Our adapted taxonomy, described below in Section 4, is built around the idea that perception in MR is influenced by factors originating in both the virtual and physical environments. The general intent of the modifications is to frame user experience as occurring
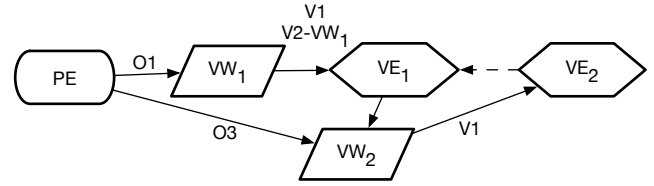

Figure 2: A graph depicting the contextual nesting of immersive environments in the see-through mode of VARI^3. Solid directional edges represent contextual relationships as the user is immersed in multiple environments. The dashed edge indicates that VE_2 is a reference to VE_1. States are designated above some edges, to be described below in Section 4.

in the context of nested immersive environments.

With the nesting scheme, we can visualize the contextual relationships between the physical environment and, potentially, multiple, virtual environments. This will allow us to understand each level at which perception is impacted. All physical and virtual environments are places where immersion occurs, and where a user should feel some degree, however slight, of subjective presence. At minimum, all interactions must take place in the *physical environment* (PE). This single environment includes all physical system hardware and all other incidental physical artifacts such as the lab floor. Within the PE, some number ($n \geq 1$) of viewing windows ($VW_1, \ldots, VW_n$) must exist and one or more ($m \geq 1$) VEs ($VE_1, \ldots, VE_m$) must also exist. These can be organized into a graph depicting the contextual nesting of environments, in accordance with the following rules:

- Each viewing window must contain a view into one or more VEs and/or back into the PE;

- In addition to existing within the context of the PE, any viewing window may also exist simultaneously in the context of any VE; and

- Any VE may reference another VE or the PE, such that all objects are mirrored and any modification to one has the same effect on the other. Note, however, that links to other graph nodes are not mirrored when referencing another VE, preventing infinite cycles.

Solid edges indicate context, such that the node pointed to is perceived in the context of the node pointed from. Dashed edges indicate a reference relationship, such that the node pointed from references the node pointed to.

It is possible for a use case to include a second physical display window with views into a VE and/or the PE. For this discussion we will assume that at least one of these views exists because it does not generally make sense to consider contextual nesting if not viewing anything through a second physical display. An exception to this may occur when comparing sources of bias between users, as described below in Section 5, though we will not explicitly consider this case. In some systems, a second display may be used simply to present interface choices (buttons, sliders, etc.) that can be used to manipulate VE_1. In these cases, we consider the second display only as a traditional input device, and thus not in terms of this discussion.

As an example, consider the see-through mode in VARI^3. For this, the VE is viewed on the CAVE displays as well as on the tablet, which can be moved flexibly though the physical and virtual environments. We can visualize the nesting as in Figure 2. The user (U_1) is physically situated in the PE. He looks through VW_1 (CAVE displays) into VE_1 and through VW_2 (tablet display) into

VE_2, which is in fact a reference to VE_1. Because VW_1 is an immersive CAVE, U_1's interactions with VW_2 also occur within the context of VE_1 as he is already immersed in that environment.

4 AN ADAPTED TAXONOMY FOR COMPUTER-MEDIATED VIRTUAL REALITY

While VARI^3 does fit within Stoev and Schmalstieg's taxonomy [13], such classification fails to acknowledge the physical aspects of the system. Implicitly, 100% virtual presence is assumed, which is unlikely to ever be the case.

We propose modifications to the existing taxonomy to emphasize that the context in VR systems with multiple heterogeneous displays is ultimately physical. Recall in the original taxonomy that there is a primary environment and it has a copy, known as the secondary environment, that can be viewed through an output window (called the viewing window, in the secondary environment). Here we make minor redefinitions and remove the equivalence between primary and secondary environments. In our framing, as described above in Section 3, all windows into secondary environments exist at least in the PE. It is these relationships between secondary environments and the PE that we will define using the taxonomy. Importantly, every physical display device exists in the context of the PE (at least) and should be considered a window into another environment (any VE) and/or back into the PE. Because graphical elements can be overlaid on a display window, it is possible that one window looks into multiple environments.

Recall the O and V states from Stoev and Schmalstieg's taxonomy. Similar states can be described for the new taxonomy. First consider the following possible states for the output windows with respect to the PE.

Case O1: Fixed in the PE.

Case O3: Mapped onto a mobile tablet in the PE.

Note that there is no Case O2, as the output windows are physical displays and thus cannot be fixed relative to the image plane. Next consider the following possible states for each viewing window into nested environments (the PE or any VE).

Case V1: Nested environment fixed relative to the PE.

Case V2-VW_n: Nested environment fixed relative to VW_n.

Case V3-U_u: Nested environment fixed relative to the primary viewpoint of user U_u.

Case V2 is noteworthy as it allows a nested environment to be displayed relative to any viewing window in the system. This classification is important for multi-user situations in which one user always tracks the view seen by another user, for example.

Case V3 allows for multi-user systems, in which one user views a display over another user's shoulder. If a nested environment is fixed relative to one of the users' viewpoints, this should be formalized.

The nested contexts of immersion and the taxonomy have been described with a focus on use cases involving a single-user. However, this scheme does facilitate visualization of bias asymmetry that may exist between users, as in Section 5. For example, it is possible for two (or more) users to use a system concurrently, with only one holding the tablet. In this case, both users may be immersed in VE_1 and the PE, but only one user (the holder of the tablet, perhaps) is immersed in the environment(s) viewed through a given viewing window. If VE_2 is seen through a viewing window that is only visible to one user, and if it is not simply a reference to VE_1, then information may be available to one user that is completely unavailable to the other. If VE_2 is a reference to VE_1, then both users may have access to the same information but different

biases exist for each user when perceiving that information. Further complexity is involved due to individual differences between users, which will affect the biases at each level of the nesting. Here we will not explicitly consider problems related to some users not viewing from the center of projection, but this can be considered another source of information asymmetry between users.

With a proposed taxonomy, together with rules for constructing a graph, we can formally describe VARI^3 as shown above in Figure 2. One display (VW_1) is the physical CAVE and it is always in Case O1 because it is always fixed in the PE. The other output window (VW_2) is always in Case O3 because it is a handheld tablet, which can be positioned freely in the PE. Most of the functions described in Section 1.1 operate in see-through mode. In these cases, VW_1 is fixed with respect to VE_1 so Cases V1 and V2-VW_1 cannot be distinguished. The state of VW_2 with respect to VE_2 is V1, because the window's view is fixed relative to the PE. Note that the primary difference from the classification above in Section 2.4 is that the CAVE is now considered to exist in physical space as another output window that displays a view into the VE.

5 TRACING THE SOURCES OF BIAS

Now that we have formally classified the VARI^3 system, we must consider the perceptual implications. Here we choose to focus on the see-through use cases. We can identify sources of bias by stepping through the graph shown in Figure 2 and considering each edge. First, biases affect judgements of the CAVE walls (VW_1) as seen from within the PE, because judgements in the physical world are not always accurate. As users consider physical relationships when perceiving virtual distances [6], for example, inaccurate physical judgements may also affect virtual judgements. Biases in viewing VE_1 through VW_1, due to display characteristics such as field of view or resolution, may affect a user's judgements in VR. Both of these sources of bias together may affect a user's overall impression of VE_1. U_1 also views VE_2, and related judgements may be biased by the physical perception (e.g., orientation and position) of the tablet (VW_2) together with biases involving the tablet's display of VE_2. Additionally, as these judgements occur in the context of U_1 being immersed in VE_1, all biases related to VE_1 may factor into the perception of VE_2.

Next consider the addition of a second user (U_2) to the system. U_2 is also physically situated in the PE and views VE_1 through VW_1. However, in contrast to U_1, U_2 is not holding the tablet and thus generally does not have a view through VW_2 into VE_2. In this case, we can see that a bias asymmetry exists between U_1 and U_2. This scenario can be further complicated, by allowing U_2 to view VE_2 through a different tablet display (VW_3), as shown in Figure 3. This arrangement still results in asymmetric bias between users.

In addition to biases being added at each graph edge, information may be gained as well. The information gain can thus also be modeled in the same way. Importantly, the intent is not to allow for precise calculations of bias and information in a given system. Indeed this is a more complex endeavor. For example, it is likely that the overall bias when concurrently using two displays may be different than the sum of the parts. Instead, this scheme is intended to provide a basis for comparing systems and to guide experiments.

6 CONCLUSION AND FUTURE WORK

In this paper, we attempt to fit our new system, in which we augment an immersive VE with a tablet display, within existing taxonomies. Such a classification fails to account for the nesting of immersive environments, including the physical one. This limits its usefulness for identifying sources of perceptual bias when augmenting a VE in which a user is already immersed.

An added window into the VE may affect a user's perception of the scene, so we propose a scheme to explicitly consider the
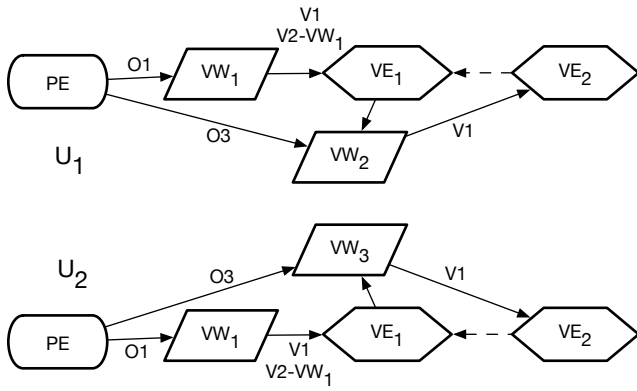
Figure 3: Graphs depicting the nesting of immersive environments in a system like VARI³, but with the addition of a second user (U_2) with a second tablet (VW_3).

contextual nesting relationships between immersive virtual environments and the physical environment. An existing taxonomy was then modified to describe through-the-lens systems with multiple heterogeneous displays in terms of these nested environments.

Acknowledging that all interaction takes place in the context of the physical environment and that even immersive virtual environments can be nested is an important first step toward identifying sources of perceptual bias. The next step is to conduct experiments which systematically evaluate user perception in VARI³. This evaluation should address questions related to the integration of information, such as distances and design attributes, as well as perspective-taking performance between the two displays. These studies should be guided by the nesting scheme and taxonomy proposed in this paper.

ACKNOWLEDGEMENTS

REFERENCES

[1] R. Aspin. An initial study into augmented inward looking exploration and navigation in CAVE-like IPT systems. In *Proceedings of IEEE Virtual Reality (IEEE VR)*, pages 245–246. IEEE, 2008.

[2] J. R. Cauchard, M. Löchtefeld, P. Irani, J. Schoening, A. Krüger, M. Fraser, and S. Subramanian. Visual separation in mobile multi-display environments. In *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, pages 451–460. ACM, 2011.

[3] D. Drascic and P. Milgram. Perceptual issues in augmented reality. In *Proceedings of SPIE*, volume 2653, pages 123–134, Bellingham, WA, 1996. SPIE.

[4] J. M. Loomis and J. M. Knapp. Visual perception of egocentric distance in real and virtual environments. In L. J. Hettinger and M. W. Haas, editors, *Virtual and Adaptive Environments: Applications, Implications, and Human Performance Issues*, pages 21–46. CRC Press, Mahwah, NJ, 2003.

[5] J. Looser, M. Billinghurst, and A. Cockburn. Through the looking glass: The use of lenses as an interface tool for augmented reality interfaces. In *Proceedings of the 2nd international conference on computer graphics and interactive techniques in Australia and south east Asia (GRAPHITE)*, pages 204–211. ACM, 2004.

[6] W. E. Marsh, J.-R. Chardonnet, and F. Merienne. Virtual distance estimation in a CAVE. In C. Freksa, B. Nebel, M. Hegarty, and T. Barkowsky, editors, *Spatial Cognition IX*, Lecture Notes in Artificial Intelligence 8684, pages 354–369, Switzerland, 2014. Springer.

[7] P. Milgram and F. Kishino. A taxonomy of mixed reality visual displays. *IEICE Transactions on Information Systems*, E77-D(12), 1994.

[8] D. Norman. *The Design of Everyday Things*. Basic Books, revised and expanded edition, 2013.

[9] D. R. Proffitt, J. Stefanucci, T. Banton, and W. Epstein. The role of effort in perceiving distance. *Psychological Science*, 14(2):106–112, 2003.

[10] U. Rashid, M. A. Nacenta, and A. Quigley. Factors influencing visual attention switch in multi-display user interfaces: A survey. In *Proceedings of the International Symposium on Pervasive Displays*, New York, NY, 2012. ACM.

[11] J. Sanneblad and L. E. Holmquist. Ubiquitous graphics: Combining hand-held and wall-size displays to interact with large images. In *Proceedings of the Working Conference on Advanced Visual Interfaces*, pages 373–377, New York, NY, 2006. ACM.

[12] M. Spindler, S. Stellmach, and R. Dachselt. Paperlens: Advanced magic lens interaction above the tabletop. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces (ITS)*, pages 69–76, New York, NY, 2009. ACM.

[13] S. L. Stoev and D. Schmalstieg. Application and taxonomy of through-the-lens techniques. In *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST)*, pages 57–64, New York, NY, 2002. ACM.

[14] Z. Szalavári and M. Gervautz. The personal interaction panel - a two-handed interface for augmented reality. In D. Fellner and L. Szirmay-Kalos, editors, *Proceedings of Eurographics*, volume 16, pages 335–346, 1997.

[15] L. Terrenghi, A. Quigley, and A. Dix. A taxonomy for and analysis of multi-person-display ecosystems. *Personal and Ubiquitous Computing*, 13(8):583–598, 2009.

[16] B. G. Witmer and M. J. Singer. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments*, 7(3):225–240, June 1998.