Yujie Tao University of Chicago yujiet@uchicago.edu Pedro Lopes University of Chicago pedrolopes@uchicago.edu



Figure 1: We explore how to minimize distractions that are detrimental to immersion in virtual reality. (a) This user's immersion was broken by a gust of wind when someone opened the door; unfortunately, these types of external stimuli are common and very hard to block (i.e., while some sounds can be minimized using noise-canceling headphones, imagine how one would block wind, temperature shifts, vibrations, smells, etc.). To tackle this challenge, (b) we explore integrating real-world distractions into the user's VR experience. With our system, the gust of wind that would otherwise distract the user is now mapped into VR: the user sees wind that sways trees, which feels more immersive and less distractive as they also feel a physical wind.

ABSTRACT

With the proliferation of consumer-level virtual reality (VR) devices, users started experiencing VR in less controlled environments, such as in social gatherings and public areas. While the current VR hardware provides an increasingly immersive experience, it ignores stimuli originating from the physical surroundings that distract users from the VR experience. To block distractions from the outside world, many users wear noise-canceling headphones. However, this is insufficient to block loud or transient sounds (e.g., drilling or hammering) and, especially, multi-modal distractions (e.g., air drafts, temperature shifts from an A/C, construction vibrations, or food smells). To tackle this, we explore a new concept, where we directly integrate the distracting stimuli from the user's physical surroundings into their virtual reality experience to enhance presence. Using our approach, an otherwise distracting wind gust can be directly mapped to the sway of trees in a VR experience that already contains trees. Using our novel approach, we demonstrate how to integrate a range of distractive stimuli into the VR experience, such as haptics (temperature, vibrations, touch), sounds, and smells. To validate our approach, we conducted three user studies

UIST '22, October 29-November 02, 2022, Bend, OR, USA

© 2022 Copyright held by the owner/author(s). Publication rights licensed to ACM. ACM ISBN 978-1-4503-9320-1/22/10...\$15.00 https://doi.org/10.1145/3526113.3545682 and a technical evaluation. First, to validate our key principle, we conducted a controlled study where participants were exposed to distractions while playing a VR game. We found that our approach improved users' sense of presence, compared to wearing noise-canceling headphones. From these results, we engineered a sensing module that detects a set of simple distractive signals (e.g., sounds, winds, and temperature shifts). We validated our hardware in a technical evaluation and in an out-of-lab study where participants played VR games in an uncontrolled environment. Moreover, to gather the perspective of VR content creators that might one day utilize a system inspired by our findings, we invited game designers to use our approach and collected their feedback and VR designs. Finally, we present design considerations for mapping distracting external stimuli and discuss ethical considerations of integrating real-world stimuli into virtual reality.

CCS CONCEPTS

• **Human-centered computing** → Human computer interaction (HCI); Interaction paradigms; Virtual reality.

KEYWORDS

Virtual Reality, Haptics, Distractions, Presence

ACM Reference Format:

Yujie Tao and Pedro Lopes. 2022. Integrating Real-World Distractions into Virtual Reality. In *The 35th Annual ACM Symposium on User Interface Software and Technology (UIST '22), October 29–November 02, 2022, Bend, OR, USA.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3526113.3545682

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

1 INTRODUCTION

Virtual reality (VR) enables users to experience being in an environment beyond where they physically are. The emergence of portable VR hardware (e.g., Oculus Quest 2, VIVE Focus 3) further allows users to experience virtual reality anywhere and anytime. However, current VR systems, while rendering immersive visual and audio experiences, are mostly ignorant of events in the user's physical environment. In contrast, a user experiencing VR constantly receives two streams of sensory information, one from VR and one from the physical surroundings. Many real-world cues (e.g., background noises, wind, smell, cold air, etc.) can intrude into or contradict the virtual experience. A break in the presence happens when the user shifts their attention away from the virtual world to the real-world sensory stream [7, 37, 44]. A study by Slater and Steed showed that "external sound" and "external touch" were rated as the top two causes that break the presence [37]. To block these distractive signals, the predominant solution relies on noise-canceling headphones [22]. However, even noise-canceling headphones can only block sounds up to a certain intensity and, more importantly, do not block multimodal distractions such as haptics or even smell.

In this work, rather than blocking or ignoring distractive signals from the outside environment, we propose *integrating* them into virtual reality to improve the sense of presence. Here, distractions are defined as external signals perceived by VR users that *do not match* the consequences of the experience they have in the virtual world. A common example would be a user playing a VR game at home, but their housemates are cooking in the kitchen, which carries out food smells. Using our technique and a VR experience that was already prepared to leverage it, we can *map* the external food smell to the arrival of a food truck in the virtual experience this allows users to feel a coherent VR experience.

To understand whether the integration of distractions enhances presence, we conducted an in-lab study where participants were exposed to controlled distractors while in a VR game. We found that our technique, including its two mapping approaches (*direct & stretch* mapping), enhanced participants' presence and decreased the perceived distraction. Based on these results, we engineered a sensing module that can be attached to VR headsets and detects a set of distractive signals (speech, engine sound, door closing sound, wind, and temperature shift). We validated it in a technical evaluation (i.e., accurately detected 77.2% of distractive signals) and in an out-of-lab study where participants played a VR game in an uncontrolled environment (i.e., observations followed the trend from the controlled study).

Moreover, to gather the perspective of VR content producers that might one day utilize a system inspired by our findings, we invited game designers to try out our approach. Finally, we present design considerations and discuss ethical considerations of integrating real-world stimuli into virtual reality.

2 WALKTHROUGH: A ROOM ESCAPE EXPERIENCE

To help readers understand the applicability of our technique, we demonstrate it in a VR "escape room" experience. In Figure 2a, the user stands in a multi-user recreational room. They use a Quest 2 VR headset with our sensing module attached to it—this hardware

Yujie Tao and Pedro Lopes

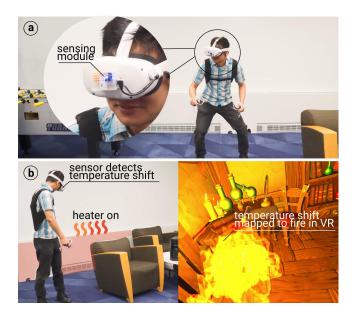


Figure 2: (a) A user is playing a VR in a recreational room with our sensing module attached to their headset to detect distractions. To prevent a break in presence, (b) our sensor detects that the room temperature changed to 27.05° C, and (c) instructs the VR to render a pre-defined effect for heat distractions (a fire breaks out in the virtual room).

can detect simple distraction events, such as sounds, temperature shifts, and wind around the user.

In the VR experience, the user must find the keys to escape the VR room. However, while they are immersed in VR, the room's air conditioning turns on to heat the room, which creates a *feelable* temperature change. Without our concept and underlying implementation, the user could feel this heat as a distraction—this heat is not consistent with the virtual experience, and **it could cause a break in presence**. Instead, our system detects the temperature shift and masks this potential source of distraction with a VR effect that "explains" it. Specifically, the VR experience renders a pre-defined effect in which a fire appears in the virtual room (Figure 2b). This exemplifies the core of our approach: rather than ignoring or blocking external stimuli that can disrupt presence, we integrate such distractive stimuli directly into the player's VR experience.

Next, the user searches for the mystery keys around the writing desk while the fire VR effect fades out (as it was pre-designed to do, it was inconsequential to the game's narrative or mechanics). However, as they approach the VR desk, another person comes into the recreational room and turns on the fan, which creates wind, as depicted in Figure 3a. Using our technique, the VR responds by triggering a pre-designed VR effect that animates the window curtain (already in the scene) to match the felt wind sensation.

Next, the player explores the bookshelf in search of the missing key. At the same time, the other occupant of the recreational room starts the coffee machine, which creates a long "motorized" kind of sound, as shown in Figure 3b. To tackle the distractions, the VR

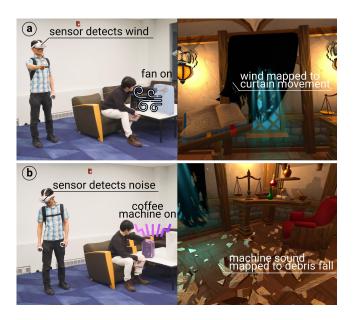


Figure 3: (a) Another occupant turns on the fan, creating a wind that could distract the VR user. Our system detects it and instructs the VR to render an effect for this distraction (a curtain moving). (b) Now, the occupant turns on a coffee machine, which creates sounds; again, our sensor detects an "engine" sound and instructs the VR to render an effect for this distraction (debris falling from the ceiling).

experience responds by triggering a pre-designed VR event that maps the sound to the fall of debris from the ceiling.

Finally, this walkthrough is meant to illustrate a subset of what our concept can provide. In fact, this VR experience demonstrates how we can *directly* map external distractions into virtual reality (e.g., a temperature change maps well to a fire, a wind maps well to a curtain moving) but also how we can *stretch* the relationship between distraction and its VR counterparts (e.g., the coffee machine sound maps to the fall of debris, which is not a *direct* mapping). We also explore how our technique can handle distractions ranging from sounds to multimodal factors such as haptics and smell.

3 OUR APPROACH: INTEGRATING SURROUNDING DISTRACTIONS INTO THE VR EXPERIENCE

The key concept behind our approach is to integrate distractions (from the user's physical surroundings) into the user's virtual reality experience to reduce breaks in immersion in VR. We define distractions as external signals *perceived* by VR users that *do not match* the consequences of the experience they have in the virtual world. Unfortunately, these distractions are very often detrimental to users' immersion.

Figure 4 depicts the workflow required to instantiate our approach on the end-user side: (a) a VR user enables "distraction integration" by selecting which stimuli they allow our system to integrate; this leaves users with full agency in controlling our system's behavior. Then, (b) the user wears a VR headset with our sensor

UIST '22, October 29-November 02, 2022, Bend, OR, USA



Figure 4: (a) A end-user selects which distractions to be integrated with their experience. (b) The sensor module detects distractive stimuli, in this case, a wind gust. (c) A pre-defined mapping design is triggered to map with the wind.

module attached to it. Our proof-of-concept prototype detects different types of stimuli, including sounds, wind, and temperature shifts. As such, (c) when a user-allowed distraction is detected, such as this wind gust, our sensing module notifies the VR scene, which causes it to render pre-designed VR effects that map this distraction (here, the trees sway with the wind gust). As we will see in Study #1 and #2, we found that integrating distractions improved presence and reduced perceived distraction.

Figure 5 depicts the workflow required to instantiate our approach on the VR creator side: (a) a VR designer creates interactions/scenes/effects to handle the distraction (here, swaying trees and wind visuals). (b) Using our simple unity mapping script, the designer drags and drops the reference of the *game object* to the stimuli they want to mask, and the script will call the designed interaction upon any detection of this stimulus (if the user enables it). Finally, (c) the designer tests their integration on their VR editor, assisted by a simple script that can simulate sensor detection results. As we will see in Study #3, we found that our recruited VR designers were able to create a wide range of experiences by integrating distractions into virtual reality.

To better illustrate the scope of our approach, Figure 6 depicts exemplary distraction sources that our approach can potentially integrate. We focus on *external* stimuli perceived by users that do not align with the VR experience. As VR headsets already block most visual stimuli from the physical world, our technique is most suited for audio, haptics, and smell. Note that Figure 6 does not present an exhaustive design space and only intends to show a subset of stimuli that our technique can potentially integrate. Specifically, we include signals that commonly exist in our environment and could/soon be detected by sensors, to provide design possibilities while taking into consideration of technical feasibility.

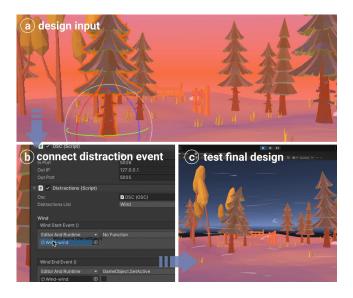


Figure 5: (a) A designer experiments on how to mask distractive wind stimuli. (b) The designer drags the object and its designed interaction into our custom mapping script to connect with sensor detection results. (c) The designer tests the final design.

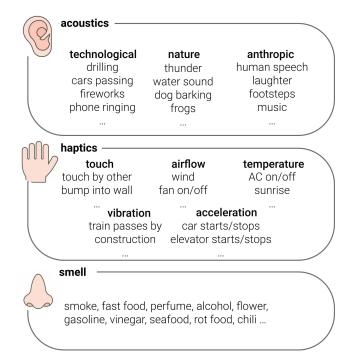


Figure 6: Exemplary stimuli that our approach can potentially integrate, which lie in acoustics, haptics, and smell.

To integrate distractions into the virtual experience, we explore two types of designs, which we call *direct* and *stretch* mapping. In a *direct* mapping, an external stimulus is mapped to a VR counterpart that directly attempts to represent the source of distraction

(e.g., physical wind mapped to VR wind effects)-this is akin to skeuomorphism [25, 47] or biomimicry [29] for traditional screenbased UI designs. Conversely, in a stretch mapping, the designer "stretches" the connection between the external stimuli and the VR counterpart, representing the source of distraction more loosely (e.g., physical wind mapped to a VR ghost that travels through the user's body)-this is akin to how metaphorical references [2] are used in screen-based UI designs to depict concepts that are not exactly the same as in the real world. The concept of stretch mapping is also inspired by prior work on tolerance of sensory mismatch in the VR [1, 15, 34]. Substitutional Reality [34], for instance, investigated passive haptic with a certain degree of visual-haptic mismatch; in our paper, we explored the discrepancy between the true source of the distraction and its mapping. The advantage of exploring two designs, rather than just *direct* mappings, is that *stretch* mappings can reduce the amount of work for VR content creators while still minimizing distraction (see Study #1).

4 BENEFITS, CONTRIBUTIONS AND LIMITATIONS

Our key contribution is an approach to minimizing distractions originating in the user's physical surroundings, which are often detrimental to the user's immersion.

Our approach has the following benefits. (1) Provides a new way to increase immersion that can be applied to a variety of VR experiences; in fact, the core of our approach is modifying the user's virtual experience, which is easier and cheaper than changing the user's physical environment or hardware. (2) Wide reach; even in its simplest form, our approach can already work with the sensors existing in today's VR headsets-this means our approach has an immediate and wide application; taking the Oculus Quest as an exemplary consumer device, we can see how its microphone can be leveraged to detect auditory distractions (e.g., loud transient sounds, wind sound), its accelerometer might provide a simplistic way to detect large ground vibrations (e.g., floor rumbling) or accelerations (e.g., while playing VR inside a car), and so forth. (3) Sensing hardware has a small footprint typically and sensing approaches are evolving rapidly; since our approach only requires hardware for detecting the distraction sources, its footprint can remain small (sensors are typically smaller than actuators); moreover, we expect new sensing methods will integrate new capabilities beyond those of our current implementation.

Our approach is not without its limitations: (1) Content creation; our approach requires the designer's input to conceptualize and create VR counterparts for each distraction they are targeting; while our study with VR designers showed that they found it easy and enjoyable to generate the mappings, it still costs time for content creators. One immediate way to minimize this is our *stretch mapping designs*, which associate a physical stimulus to a virtual experience that only *coarsely* matches the consequences of this stimulus; we also recommend designers design one-to-many mapping across different types of stimuli to help mitigate the effort, as described in design considerations; also, another promising direction is to automate the creation of these VR effects. (2) Handling useful interruptions; while many distractions are indeed detrimental to the user's immersion (e.g., noises or wind), there

are also stimuli that carry useful information to the user, such as a friend calling the user's name for attention. Currently, our system does not implement a way to distinguish these (i.e., coarse classification, no distinction between "talking sounds" to "someone is calling me"). We expect that as recognition systems advance, some of these features are possible and highly desirable to implement for our approach. Moreover, to mitigate the impact of this limitation, we implemented an interface that enables users to select the type of distractions they allow their VR experience to integrate. (3) **Creating unnecessary game events**; our approach relies on the detection of distractions; thus, misdetection could cause unnecessary game events, confusing the VR experience; one way to tackle these is to create designs that are sensible in VR and will not feel out of place even in the absence of external stimuli; we expect that advances in robust sensing will further minimize this limitation.

5 RELATED WORK

Our work is inspired by prior research on presence in virtual reality and synchronization between virtual experiences and the physical environment that surrounds the user.

5.1 Presence in virtual reality & break in presence

Virtual reality hardware creates a sense of "being there" by rendering sensory stimuli simulating what we perceive from the physical environment [14]. The VR interface's capacity to deliver a vivid experience that removes the user from physical reality is often referred to as immersion, while presence denotes a state of consciousness, the psychological sense of being in the virtual environment [35, 38]. When a user is experiencing virtual reality, their body constantly receives two streams of sensory information: one from the virtual environment displayed by the virtual reality hardware, and the other from the real-world environment the user is physically in [36]. Thus, a **break in the presence** happens when the user shifts their attention away from the virtual world to the real-world sensory stream [7, 37, 44]. Slater et al. [37] further denoted the reasons why a break in presence happens to both external factors (sensory information from the physical world intrudes or contradicts that of the virtual world) and internal factors (something wrong with the virtual experience itself). Oh et al., [26], for example, explored how external stimuli (such as a ringing cell phone) can have an impact on cognitive, affective valence, and interpersonal outcomes in virtual social interaction.

5.2 Blending real-world environment into virtual reality

To blend the real-world environment into virtual reality, one line of work focused on leveraging physical objects to enhance the sensory experience in VR. Early work in the passive haptics [15, 16] investigated the impact of registering physical objects to virtual objects and receiving feedback from physical interaction. Cheng et al. [3] proposed a set of haptic proxies using a general passive prop for VR experience. *Sensory VR* [12] extends the exploration and bridges eating and smelling to enhance the virtual experience. Another line of work focused on visually mapping the physical environment to the VR environment. To procedurally generate virtual reality

environments based on physical spaces, Sra et al. [39] and Shapira et al. [33] proposed systems that scan and reconstruct users' surroundings. Substitutional Reality [34], studied how physical objects and architectural features can be substituted with VR counterparts, even if these have a certain level of mismatch. These systems allow daily objects to serve as haptic props. RealityCheck [13], offers another take by combining 3D reconstruction of the real world with the virtual environment. Visually mapping physical space into virtual reality also allows for free VR locomotion. VRoamer [4] enables users to walk in unseen physical spaces for which the system generates a virtual scene on-the-fly. Researchers also explored a VR tracking system that allows real walking in the outdoor environment [46]. Other work presented solutions [20, 21, 41] to design virtual world appearance into real-world geometry. While both lines of work align virtual and physical experiences, they assume our environment is synchronized with the virtual experience to enhance presence. For instance, passive haptics requires environments to possess or be instrumented with props that match the virtual environment. Current work ignores the ad-hoc environmental stimuli that contradict/mismatch with the VR experience. Those events are often detrimental to VR immersion and break the sense of presence, which is what our work focuses on.

5.3 Handling real-world interruption in virtual reality

To handle interruption from the physical world, another line of research tried to design notification systems that allow users to be informed of real-world events while in virtual reality. NotifiVR [9], for instance, explored notification design for physical, digital, and temporal events and interruptions. Similarly, Rzayev et.al [31]. and Hsieh et.al. [17] explored different design patterns and placements of notification messages to connect VR users to real-world events. Researchers also investigated how to handle bystander awareness and interruption in VR [8, 10, 27, 48]. Beyond notifying VR users through digital notifications, there is a growing amount of work that proposed new headset hardware to enable users to see both physical and virtual environments without taking off the device [5, 6, 42]. While past works provide exciting solutions to real-world interaction while keeping the users in VR, they focus on signals intended to interrupt the VR experience (e.g., a bystander trying to talk to the VR user) which have precisely the goal of bringing the users out of VR. However, this is not representative of all the events around a VR user. There are many more events happening in the environment (including many that do not have the purpose of interrupting the experience but do so as a side-effect). Those distractions are well under-explored in current systems, and we propose integrating them into virtual reality.

6 USER STUDY #1: INTEGRATING REAL-WORLD DISTRACTIONS IMPROVES PRESENCE IN VR

Our first study focused on validating if integrating real-world distractions directly into the user's virtual reality experience improves their sense of presence in VR. To realize this, we invited participants to our lab to play a custom-made VR game while we triggered prescripted external distractions, ranging from drilling sounds, and

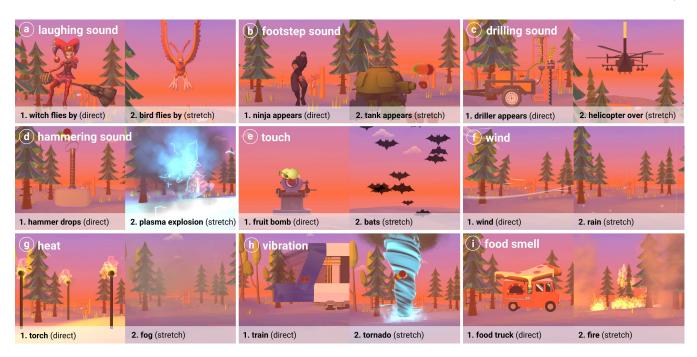


Figure 7: Study #1 mapping design used in custom fruit ninja game. It involves four sound stimuli: (a) laughing sound, (b) footstep sound, (c) drilling sound, (d) hammering sound, and five multimodal stimuli: (e) touch, (f) wind, (g) heat, (h) vibration and (i) food smell. For each distraction, we designed both a *direct* and *stretch* mapping that attempted to integrate this distraction into VR.

wind, to food smells. We adopted a within-group study design, in which participants experienced *no mapping* (baseline), *direct mapping*, and *stretch mapping* of these distracting stimuli into the VR experience. Our hypotheses were: **(H1)** our technique would improve the sense of presence compared to the baseline; **(H2)** our technique would reduce the amount of distraction felt while in VR compared to the baseline; **(H3)** that *stretch mapping* would be less effective than *direct mapping* in terms of presence, but still improve the sense of presence compared to the baseline.

This study was approved by our Institutional Review Board (IRB22-0265).

6.1 Interface conditions

In this study, participants played a custom-made VR game inspired by the popular *fruit ninja* game. Participants experienced three interface conditions in a counterbalanced order: *baseline* (no mapping) vs. our two designs, *direct* mapping, and *stretch* mapping. In the *baseline* condition, participants wore noise-canceling headphones. In the *direct* condition, each external stimulus was mapped to a VR counterpart that directly attempted to represent the source of distraction (e.g., a hammering sound around the participant was masked as a VR hammer dropping from the sky). In the *stretch* condition, we "stretched" the connection between distractions and their VR counterparts, attempting to represent the source of distraction more loosely (e.g., a hammering sound around the participant was masked as a VR explosion).

6.2 Stimuli (distractions)

Each participant received the same set of distractions for each condition: four sounds (laughing, footstep, drilling, hammering) and five multimodal events (touch on participants' back, wind, heat, ground vibration, food smell). These stimuli were selected to represent a wide spectrum of stimuli that our approach can potentially integrate. Figure 7 shows the full list of stimuli and their VR counterparts (both *direct* and *stretch* mapping) in our custommade fruit ninja VR game.

As this study was meant to evaluate the concept but not its technical feasibility (e.g., tracking accuracy, real-world deployment, etc.), the real-world stimuli and the VR counterparts were artificially triggered by the experimenter at fixed timings. Despite this, we still included technical factors that an actual system will always encounter, such as end-to-end latency or not knowing the directionality of the stimuli (e.g., which direction a sound came from). Our triggering of VR counterparts was achieved via Open Sound Control messages; thus, Wi-Fi network latency was already included in our study, which was approximately 176.5ms. Beyond that, we artificially added extra time lag between the stimuli and the creation of VR mapping to simulate signal detection latency: 53ms [30] was added to non-speech audio distractions (footstep, drilling hammering); 140ms [19] was added speech audio and vibration stimuli; for touch, heat, and wind we added 70ms, 80ms, and 300ms latency; smells tend to display the highest detection latency, so we added 5s, considering the existing odor detection systems-these latencies were determined based on prior work in touch [11], temperature [23], wind [40], and odor [45] sensing that report their

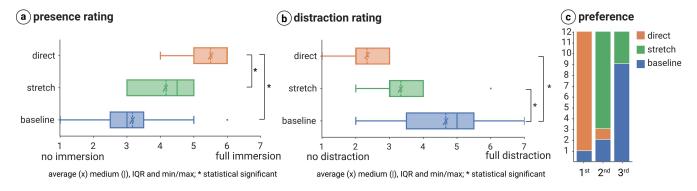


Figure 8: User study #1 result. (a) presence rating and (b) distraction rating for all three conditions (no mapping, *direct* mapping, and *stretch* mapping of distraction into VR); (c) user preference ranking of three conditions.

detection latency. We believe that future detection systems can only be faster. Furthermore, we considered the case where the directionality of the distraction source is unknown as many sensors do not detect this. Thus, we intentionally created all sound stimuli from the participants' right side while showing VR counterparts at the front. This allows us to simulate a directionality mismatch and understand if it is necessary to design all VR counterparts tightly at the same spatial location as the physical stimuli.

6.3 Apparatus

Participants wore *Bose QuietComfort 35* noise-canceling headphones in all conditions, which is today's best approach to creating a VR experience free of auditory distractions. All audio distractions were played from a speaker at 70 dB, from the participant's right side. The experimenter created a touch stimulus by touching the participants' back with a plush toy and delivered wind by waving a plastic plate. The ground vibration was achieved by playing low-frequency sound through a subwoofer attached to the wooden platform, on which participants stood for all trials. The heat was created via a heat lamp. Lastly, the food smell was manually dispersed by opening a sealed *Tupperware* with warm fried potatoes.

6.4 Study procedure

We started the study with a 5-minute training session where participants got familiarized with VR. Then, the participants experienced three interface conditions in a counterbalanced order, each lasting 10 minutes. All the distractive stimuli were presented one after another spaced by 1 minute. After each trial, we asked the participants to rate their presence and distraction level during the last session. This presence question was derived from the *Presence Questionnaire* [43]: "Please rate your sense of being in the virtual environment, on a scale of 1 to 7, where 7 represents your normal experience of being in a place." We then conducted a short interview, in which we asked participants to elaborate on their experiences. To avoid sequence effects, we presented all distractions in a randomized order. We also randomized the time gap between each distraction, ranging from 56 to 64 seconds.

6.5 Participants

We recruited 12 participants (one identified as non-binary, three as female, and eight as males), with an average age of 23.5 years old (SD = 1.5). Participants received a compensation of \$20 for their time.

6.6 Results

Figure 8 depicts the results of our study. Overall, we found that *direct* mapping improved the sense of presence while both *direct* and *stretch* mapping lowered the perceived distraction level from the physical environment compared to our *baseline* (noise-canceling headphones). Also, participants preferred both *direct* and *stretch* mappings to our *baseline* and rated *direct* mapping as the most preferred interface, as depicted in Figure 8c.

We analyzed our data using one-way ANOVA and found a significant difference among the three interface conditions for both presence level (p < 0.00005) and distraction level (p < 0.00005). Thus, pair-wise Tukey multiple comparisons were conducted. The Bonferroni-corrected alpha level is 0.0167. We present these results grouped by our three hypotheses:

H1 (presence). We found a significant difference between the perceived presence in *direct* and *baseline* (p = 0) and between *direct* and *stretch* (p = 0.01), which we discuss in H3 in detail. The results showed that *direct* mapping of distraction into VR improves the sense of presence (M = 5.5, SD = 0.8) compared to the baseline (M = 3.2, SD = 1.3). While there is no statistical difference in presence rating found between the *stretch* mapping (M = 4.2, SD = 0.9) and *baseline* condition when participants wearing a noise-canceling headset (p = 0.06), most participants (9 out of 12) preferred stretch mapping of distraction into VR to no mapping (*baseline*). These results suggest that our approach—especially *direct* mapping—improved the sense of presence when compared to the *baseline*, which confirms H1.

H2 (distraction). We found a significant difference between the perceived distraction level in *direct* and *baseline* (p = 0) and between *stretch* and *baseline* (p = 0.015). Results showed that both *direct* mapping (M = 2.3, SD = 0.6) and *stretch* mapping (M = 3.3, SD = 1.0) lowered the distraction level compared to only wearing noise-canceling headphones (M = 4.7, SD = 1.4). Taken together, these results confirm H2.

H3 (direct > stretch). Furthermore, we had hypothesized that stretch mappings would be less effective than direct mapping in terms of presence, but still improve the sense of presence when compared to the *baseline*. Indeed, we found (p = 0.01) that *direct* mapping improved the VR presence compared to *stretch* mapping. However, no statistical difference in distraction rating was found between the *direct* and *stretch* mapping (p = 0.08). Still, while *stretch* mappings were generally perceived as less effective than direct mapping, they were found to minimize distractions when compared to the baseline and were preferred over baseline (as discussed in H1 and H2). Taken together, these results suggest that stretch mappings are less effective than *direct* mappings, but still more effective than baseline (H3). Moreover, one advantage of stretch mapping is that designers can minimize their effort by creating mappings loosely related to the source of distractive stimuli and reusing the mapping designs.

Participants' preference. Figure 8c depicts the result of the preferred interface conditions as ranked by our participants. 11 out of 12 participants reported *direct* mapping as their first preferred interface condition (their top choice) and the remaining participant preferred the *baseline*. Subsequently, when asked which they preferred as their second top choice, nine out of 12 participants rated *stretch* mapping, one preferred *direct* and two preferred the baseline. Finally, eight participants chose the *baseline* as the least preferred condition and three chose *stretch*.

Finally, since our study adopted a within-group study design, we analyzed if the rating was impacted by condition order. No statistically significant interaction was found between rating and order of conditions in both presence (p = 0.7) and distraction rating (p = 0.4).

6.7 Qualitative feedback

Distractions in the baseline (noise-canceling headphones). When asked about participants' experience in VR while only wearing noise-canceling headphones (baseline), all participants mentioned that they were distracted by stimuli coming from the physical environment. P10, for instance, said: "I was playing and distraction started happening ... I swear I can smell greasy food, which made me hungry". They then elaborated on other distractions they felt during the session, such as "floor rumbled", "something touched on my back", "warmth", and "laughing sound"; they stated that these distractions took them out of VR. P6 further added that they felt stimuli that were "completely from the physical environment" and described them as "very weird (...) not part of the game".

In the baseline condition, all 12 participants reported that they felt the wind, ground vibration, and laughing sound during the session. Most of them felt heat (11 out of 12), touch on the back (9 out of 12), and food smell (8 out of 12). Drilling sound (P5, P6, P7, P8, P12), footstep sound (P1, P7, P9, P11), and hammering sound (P7) were less reported. Within 12 participants, only P5 stated the *baseline* as their most preferred interface. While P5 agreed that they felt stimuli distracting them from completing tasks in the VR game, the appearance of those stimuli, without knowing where they came from, reminded them about horror games that they really enjoyed.

Directly integrating distraction into VR. During the session with direct integration of distraction into the virtual experience,

participants reported a higher level of presence in VR and a lower distraction level. For instance, P4 reported that they saw graphics representing the wind while feeling the airflow and thought of them as "quite immersive". Similarly, P1 stated "[distractions] added to the experience as there were something happening in the game related to them; they were not distractions anymore and enhanced the experience rather". P3 and P10 added the distractions were "wellaligned" (P3), and "felt connected" (P10) with things happening in the game. P7 further elaborated that "what I felt was a result of this game, instead of things out to my control. It was interesting to see a pizza car going in front of you, and you actually smelled some food; so you feel like oh it's real!".

When designing the VR counterparts, we intentionally applied incongruent directionality: the direction of physical stimuli (e.g., the direction of the physical wind, sounds, etc.) and the direction of the VR counterpart (e.g., virtual wind) were purposefully not aligned. Only 3 out of 12 participants noted this. P7, P11, and P12 noticed this and suggested that the experience could be even more immersive if the direction of stimuli were "in tune with what I was seeing" (P11). P7, for instance, said the VR driller came from left to right but the sound came from the right side. "However, without the audio direction synchrony is still much better than no mapping," P11 added.

Stretching the mapping between a distraction and its VR counterpart. In the condition where we stretched the mapping between real-world stimuli and their VR counterparts, most participants found it less immersive than *direct* mapping, yet they still rated it higher than the baseline condition. P6, for example, reported that they saw bats around them while felt hitting on the back, they described those mappings as "still make some sense. . .but the brain needs some time to process it; in the first time [*direct* mapping], everything fit together very nicely, and this time is loosely matched". P8 also added that "it was not [a] 1-1 mapping of what I was feeling versus what I expected from the game world" and thus "made me think about the connection between the visual cues and sensation."

Upon rating *stretch* mapping to be preferable to only wearing noise-canceling headphones (*baseline*), P10 explained that while in the *direct* condition, mapping between sensory input more closely mirrored what happened in the inside VR experience, not-perfect mapping is still better than no mapping. For the 3 participants (P4, P5, P8) who rated stretch condition to be the least preferred, they explained that some environmental factors made them "more sensitive to notice the mismatch" (P8), and "[*baseline*] condition and [*stretch*] were distracting in different ways" (P4). "In the third [*baseline*] round, I know there was something happening in the real world; it was almost easier to ignore it," said P4. As in *direct* mapping, while some participants (P7, P8, P11, P12) reported stimulimapping directionality asynchrony issues, *stretch* mapping was still preferred over the *baseline* condition.

7 IMPLEMENTATION

Building on the result of our first user study, we engineered a prototype system, depicted in Figure 9, comprised of: (1) a sensor module that detects a set of distractive stimuli, (2) a user interface where users can select the distractions to be integrated into their VR experience, and (3) a simple Unity3D script that allows designers



Figure 9: Our hardware and software implementation; (a) full hardware system, (b) electronic components detail of our sensing module, and (c) our software mapping script to connect design with the sensor detection result.

to map their designs to the sensor detection result. To help readers replicate our design, we now provide the necessary technical details, and to accelerate replication, we provide all the source code, 3D files, firmware, and schematics of our implementation¹.

Hardware. Figure 9b depicts our simple prototype sensing module, which detects a limited set of distractive stimuli, including sounds (person talking, engine, door closing), wind gusts or temperature changes; these signals represent only a small subset of events that our technique could potentially integrate, as the goal of our hardware is only to demonstrate a proof-of-concept implementation of our approach and not to be exhaustive. We believe that future systems can integrate more sensing methods as well as refine sensing beyond our current implementation, to handle more distractions ranging such as contact with the user's body [24, 32] or smell [28, 45]. For sound detection, we run Ubicoustics [18] in a laptop that samples the microphone (at 44.1Khz, 16 bit), and retrieves any of the following detected labels ("door", "speech", "engine"). Our sensor module also contains an Arduino Nano microcontroller that detects temperature shifts by digital-sampling a temperature sensor (DS18B20) and detects air drafts by an analogsampling anemometer (Modern Device Wind Sensor Rev C). Our microcontroller is connected via USB to the aforementioned laptop, which users can wear in their backpack. Any wind over 10Mph or a change in temperature by 0.05°C is communicated from the Arduino to the laptop. Finally, the laptop communicates all distractions to the Oculus headset, in the form of messages using the Open Sound

Control (OSC) protocol over Wi-Fi; the VR experiences respond to these by activating their pre-designed mappings.

Software. On the software side, we implemented a simple VR user interface menu that lists all the potential stimuli detected by our system. Using this UI, end-users select which stimuli they allow to integrate, prior to the start of the VR. Considering the variety of physical environments the users might be in, such design enables the user to customize their VR experience and gives users the agency to decide which distractions are suitable. We also implemented a simple mapping Unity3D script that allows the VR designers to connect the VR effects/animations/scenes to each distractive stimuli that our sensors detect. As depicted in Figure 9c, the designer can drag and drop Unity3D object references into the UI our script exposes (using Unity3D's inspector). Moreover, they can specify which functions to trigger when each distraction starts/ends. Any events they add are therefore connected to handlers in the backend, and designers' functions will be called whenever the corresponding OSC messages arrive (and if the user enabled this type of distraction integration).

8 STUDY #2 TECHNICAL EVALUATION & OUT-OF-LAB USER STUDY

Equipped with our sensing module, we conducted two additional evaluations: (1) a technical evaluation to measure its accuracy in detecting stimuli in the user's surroundings; and (2) an out-of-lab user study, which allowed participants to experience our system without the controlled laboratory settings and scripted stimuli that we used in our first study. The latter user study was approved by our Institutional Review Board (IRB22-0265).

The summary of our findings. The results are as follows: (1) in our technical evaluation, we found that our proof-of-concept system has an average detection accuracy of 77.2% across five stimuli, and (2) in our out-of-lab study, we observed that participants' reactions aligned with our findings from Study #1 (in-lab study).

8.1 Technical evaluation

Evaluation goal. Measure the detection accuracy of our hardware prototype in uncontrolled environments.

Apparatus & locations. Specifically, we recruited a participant to wear a VR headset, attached to our sensor module, and stand in three locations: (1) an entrance hallway of a university cafe, (2) outdoors, next to a bus station, and (3) an open antechamber in a residential building. These places were chosen to cover a variety of locations where people are or might one day be experiencing VR. Experimenters did not interfere or create stimuli in these locations and just safeguarded the safety of the participant as they collected the data. All stimuli occurred naturally in these locations, such as when people passed by, or weather conditions changed.

Task. Our participant stood at the three chosen locations, each for 10 minutes. From their head-mounted display, they were presented with a list of stimuli that our sensor module detects and were asked to select any stimuli that they felt. This allowed us to compare to relevant human-level ground truth, i.e., a participant in VR. Note that the participant did not wear noise-canceling headphones for this task.

 $^{^1\}mathrm{http://lab.plopes.org/\#vr-distraction}$ (all source code, VR scenes and hardware schematics).

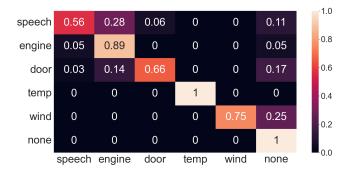


Figure 10: Confusion matrix for accuracy of our sensor module (all three locations).

Results. In total, the participant reported 29 "door" events, 18 "speech" events, 16 "wind" events, 4 "temp change" events, and 19"engine" events across all three locations. The intervals with no checked stimuli were marked as "none", which were reported 57 times. Figure 10 depicts our system's accuracy in identifying the events that the VR participant was able to perceive, across all three locations. We plotted the accuracy of our system in identifying distractions in the confusion matrix. We found an average accuracy of 77.2% across all distractions in all locations. Two results stood out. First, speech sound was the most misclassified stimulus with an accuracy of 56%. Secondly, the door closing sound was also sometimes misclassified, at an accuracy of 66%. Finally, while we can expect that advances in sensing will only push this average accuracy of 77.2% upwards, we must also note that any correct classification will result in an improvement in immersion (see Study #1 results) while missing detecting an event just default to the baseline that current VR users grew accustomed to; in other words, today's users know that there are distractions outside of VR. As such, to better understand the impact of our system, when deployed in real-time and outside the lab, we conducted another evaluation, which we present next.

Limitation. Finally, we encourage the reader to also consider the limitations of our technical evaluation: in particular, its limited sample size. We recruited a single participant (from outside of the research team) for detecting and labeling events in their surroundings, which we used as ground-truth labels.

8.2 Study #2: out-of-lab user study

Study goal. The goal of our out-of-lab user study was to explore the concept of distraction integration when deployed in real-time in an uncontrolled environment, where the distractive stimuli occurred spontaneously.

Interface conditions. In this study, participants played the VR game from Study #1. They experienced two interface conditions in a counterbalanced order: *baseline* (no mapping) and *experimental* (our distraction mapping). Since Study #1 validated that both *direct* and *stretch* mapping improved presence and lowered distraction levels, this study used a mix of both mapping strategies: wind mapped to VR tree sways, temperature mapped to VR snow, engine sound

Yujie Tao and Pedro Lopes



Figure 11: Out-of-lab study: (a) a car passing by, (b) a person talking on the phone, and (c) people with a cart passing by.

mapped to a VR helicopter, speech sounds mapped to a VR ninja passing by, and door sound mapped to VR explosion.

Additional design considerations. In less controlled environments, such as the outdoor environments in which we conducted this study, distractions might occur often, even simultaneous or overlapping. While the most straightforward approach is to trigger all the VR counterparts of any detected distraction, the rapid amount of changing sensory information in VR might overwhelm the user-this in turn would be counterproductive as the distraction mappings themselves could become a distraction. As such, in this study, we made several design decisions (guided by preliminary pilot tests) to strike a better balance on when to trigger a distraction: (1) after a sound distraction was detected and its mapping was rendered, our system disregarded incoming sound distractions for 20s; (2) similarly, once a multimodal distraction is masked, VR disregarded incoming multimodal distractions until 10s after; moreover, for either multimodal distraction (wind, temperature change), the VR randomly chose one of the two multimodal mappings to render when detected; (3) when a sound distraction and multimodal distraction happened in sequence, the mappings were rendered normally based on detection order (no wait time as in the previous two cases).

Apparatus. Participants stood at the entrance of a residential building (one of the spaces used in our technical evaluation). Our sensor module, described in *Implementation*, was attached to the participants' VR headset. Pre-designed VR counterparts were activated when the sensor detected the stimuli. As shown in Figure 11 participants encountered a wide variety of distractions while playing VR in this location (e.g., people speaking, cars, wind, etc.).

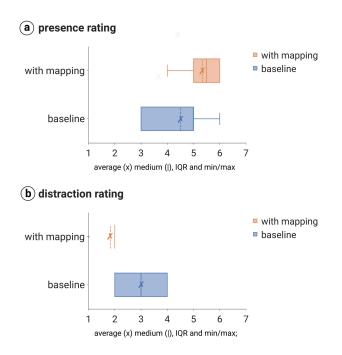


Figure 12: User study #2 result. (a) presence rating and (b) distraction rating.

Procedure. Study #2 followed a similar procedure to Study #1, i.e., participants experienced two conditions: *baseline* vs. *mapping* distractions in counterbalanced order, with each trial lasting 10 minutes and followed by a short interview.

Participants. We recruited six participants (two identified as females, four identified as males), with an average of 23.3 years old (SD = 2.4). None of the participants participated in Study #1. They received a compensation of \$20 for their time.

Result. Figure 12 depicts the key findings from Study #2. Considering the relatively small sample size (N = 6), we opted to discuss the measured data without invoking statistical tests. We observed an average sense of presence of 5.3 (SD = 0.7) in the *mapping* condition, compared to an average of 4.5 (SD = 1.1) in the *baseline*. Moreover, we observed an average distraction level of 1.8 (SD = 0.4), compared to an average of 3.0 (SD = 0.8) in the *baseline*. Both ratings appear aligned with the findings from our controlled experiment (Study #1).

Participants' preference. All participants preferred having distraction *mapping* over the *baseline*.

Qualitative feedback. All participants reported that they noticed some distractions from the physical environment when playing VR in this uncontrolled environment. The top mentioned distractions were the sounds of "cars passing by" (P1, P2, P4, P5, P6), sounds of "people talking" (P1, P2, P3, P4, P5), and wind (P1, P2, P4, P5, P6). Moreover, half of the participants reported feeling a temperature change (P2, P4, P5). When asked how these stimuli influenced their experience in the baseline condition, P6, for instance, stated that "while the visual was immersive and the elements were in harmony with each other, noises brought the presence level down." P2 added that the stimuli "distracted me [in] the short run," but

UIST '22, October 29-November 02, 2022, Bend, OR, USA

they also stated that they "can come back to the game very quickly." When asked about our mappings, P1 stated that "the wind in VR match with what I felt physically" and "those wind became not distractive that much to bring me to think about the physical environment." P2, P4, and P5 also reported that they established a link between environmental factors in VR (snow, wind) and how their bodies felt physically. For instance, P2 stated, "the added effects explain some distraction from the physical environment, making me more immersed." When considering other stimuli in the physical environment, P3 stated that "seeing people passing by [in VR, one of our VR effects] and feeling people passing by [hearing them move around the environment] made it part of VR." P2 and P5 also reported relating people walking by/talking in the physical environment to the virtual avatars in VR. P5 further noticed that "the sound of [real] cars mimics the sound of VR helicopter." They added that "some distractions from the physical environment can be regarded as effects in VR [because of the appearance of VR counterparts], which makes me more immersed", while in the baseline condition, they felt the stimuli "completely from the physical environment."

Finally, among our six participants, only one participant (P4) rated the presence level of baseline higher than the experimental condition. P4 elaborated on this by stating that even though they found "VR factor better corresponding with the physical environment" in the experimental condition, in the baseline condition, "the game mode has more predictability that I can focus on." Similarly, among our six participants, only one participant (P3) rated no difference between distraction levels between baseline and experimental conditions and further explained this rating by stating, "I am the kind of person that focuses a lot on the game".

9 STUDY #3: HOW WILL CONTENT BE CREATED? EXPERT EVALUATION BY VR DESIGNERS

While our first two studies focused on how distraction integration impacts end-users' sense of presence and awareness of surrounding distractions, in this third study, we turn our attention to the content creators-VR designers and developers who might one day interface with our technique. We conducted a design study with three game designers/developers to (1) understand their attitude toward the concept of integrating real-world distraction into virtual reality, (2) observe their design process, and (3) collect their designs and their professional suggestions. Specifically, to make the designs more comparable across participants and possible to analyze, we provided participants with a pre-built VR experience and asked them to design for six types of distractions, ranging from sound stimuli (i.e., long motorized sound) to multimodal distractions such as temperature change and smoke smell. A semi-structured interview was conducted after the design period to understand their design process and qualitative feedback.

This user study was approved by our Institutional Review Board (IRB22-0265)

9.1 Participants

We recruited three VR designers/developers (two identified as females, one as male). All of them had at least two years of VRspecific game development experience and received training from

Yujie Tao and Pedro Lopes

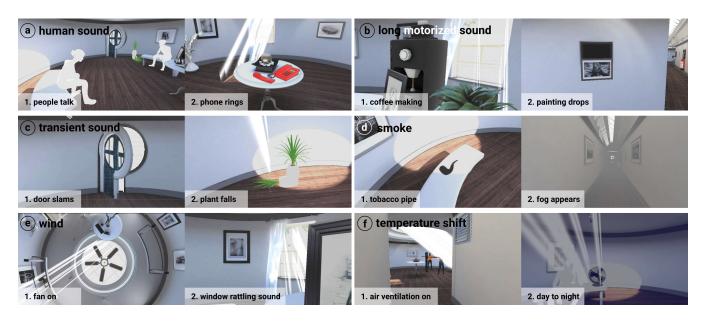


Figure 13: Selected VR distraction mappings from two participants in Study #3 (depicted with participants' written consent).

university-level game design programs. They are adept at using game engines such as Unity 3D and other design toolsets. One of our participants (P1) has worked in a major video game company and was involved in the release of several mainstream games. The participants were compensated with \$70 for their time.

9.2 Procedure

This study was comprised of three phases: (1) an introduction to the VR materials; (2) a 7-day period in which participants worked on designs on their own time; and (3) a semi-structured interview to understand design choices and collect suggestions.

Task. We provided participants with a Unity3D project that contained a VR museum experience. It included the scene, together with a unity script that received data from our sensor module (or from the event simulation script). Participants were asked to design experiences that masked six types of distractions: (1) long mechanical sounds (e.g., drilling, vehicle's motor running, helicopter flying by, etc.), (2) transient sounds (e.g., hammering, object fall, etc.), (3) human sounds (e.g., crowd or person talking, laughing, etc.), (4) wind gust (e.g., fan, outdoor wind gust, etc.), (5) change in temperature (e.g., walking from indoor to outdoor, AC on/off), and (6) smoky smell (e.g., campfire, cigarette, BBQ). For each stimulus, participants were asked to design two versions. They were not introduced to our concepts of direct or stretch mappings. Participants were told that they could change/remove/add anything into the scene, except for the museum paintings, to ensure the goal of the VR experience was respected. To allow participants to test their design together with the physical stimuli, we provided them with sound samples and a scratch-and-sniff sticker that contained campfire smells. Participants were encouraged to go outdoors to test the wind and temperature shifts.

9.3 Design outcomes & qualitative feedback

In total, our participants created 36 designs to integrate distractions into the virtual museum experience. Figure 13 and Figure 14 depict a selection of these design outcomes. We analyzed the designs and the transcribed interviews and clustered participants' results into six topics (1) mapping strategy, (2) considerations for the VR narrative, (3) designing around attention, (4) reversing cause & effect, (5) reusing mappings, and (6) non-visual mapping.

1. Mapping strategy (direct vs. stretch). We found an almost even split when categorizing the participants' designs by direct or stretch mappings. 19 out of 36 designs depicted direct mappings, in which the external stimulus was mapped to a VR counterpart that directly attempted to represent the source of distraction. For instance, P1 designed an effect in which a bystander avatar appears surrounding the VR user to mask the sound of people speaking, shown in Figure 13 (a, 1). Conversely, 17 out of 36 designs depicted stretch mappings, in which participants "stretched" the connection between distractions and their VR counterparts that more loosely attempted to represent the source of distraction. For example, P3 created a VR effect that makes a light flicker and switch off, to mask any transient sound distractions, as shown in Figure 14 (a, 2). We also zoomed into each category of distraction (e.g., sounds vs. multimodal) to examine how participants made use of different strategies. We observed that participants created more stretch mappings for auditory distractions. 12 out of 18 auditory distractions designs adopted stretch mapping strategies, while only 5 out of 18 for multimodal distraction designs. This is intuitive in that the VR headset already masks the source of the sound, so our participants took advantage of this to stretch the VR counterparts for more "ambiguous" sounds (e.g., a motorized sound can be created by a wide range of sources).

2. Fit distractions to narrative vs. alter the narrative to fit distractions. We observed that two (of three designers) kept the



Figure 14: Selected VR distraction mappings from another participant in Study #3 (depicted with participant's written consent).

original narrative of the museum experience as intact as possible, as shown in Figure 13. They designed VR counterparts that aligned with the original atmosphere of the museum, such as the "coffee machine" in Figure 13 (b, 1) that masked motorized sounds, the "plant pot that fell" in Figure 13 (c, 2) that masked transient sound, or the "ceiling fan" in Figure 13 (e, 1) that masked wind. Conversely, P3 re-designed the museum as an experience in the dark, as shown in Figure 14—P3 explained this by stating: "a lot of the stimuli do not necessarily fit with the [original] museum environment... horror uses a lot more definitive select sound that helped to create the atmosphere". For instance, P3 used this "horror" version of the museum to add surprising effects that masked the distractions, such as glitches in the TV from Figure 14 (a, 1) to mask transient sounds and making blood drop from the ceiling in Figure 14 (b, 2) to mask speech.

3. Designing with attention. Among all the designs, we observed a split between distraction integrations that try to rapidly shift the user's attention to make sense of the distraction and those that stay in the ambiance for the user to notice. 17 out of 36 designs were created to shift the user's attention, typically by showing the VR effect clearly in front of the user's viewpoint or drawing attention to it using extra sounds, forcefully redirecting the viewpoint, etc. For instance, P1 and P2 masked speech sound using bystander avatars that walk around the VR user, which users could hardly miss. "People are more sensitive to human sound and speech. We are tuned to pay attention to that," said P2. Conversely, 19 of 36 designs provided more subtle cues that were more ambient. For instance, to mask the campfire/cigarette smell, one of the P2's designs was to render a tobacco pipe on the museum bench, shown in Figure 13 (d, 1). P2 elaborated on this by stating: "integration should not be a distraction; there are chances that players didn't even catch the smell." P3 also emphasized the "ambiance" of some of her designs by stating: "I am not trying to draw the [user's] attention at all; with subtleness, you can integrate things with a lot more fluidity in it."

4. Reusing mappings. Without any instructions to do so, we observed eight instances of participants using the same VR counterpart to mask multiple distractions: (1-2) a "ceiling fan" mapped to the wind and temperature shifts; (3-4) an "air conditioner" mapped to temperature shifts and motorized sounds, (5-6) "fog" mapped to

temperature shifts and smoke smells; and (7-8) a "fireplace" mapped to temperature shifts and smokey smells. Regarding the decision rationale, P2 explained that "I am not gonna get every possible distraction, so being able to adapt and be flexible is important (...) you don't want to find yourself in a situation where you create an effect per distraction."

5. Non-visual mappings. We observed two participants (P1 and P2) that created VR effects (to mask distractions) using sound alone. For example, P1 masked a wind using the sound of a fan, as depicted in Figure 13 (e, 1). Similarly, P2 experimented with using only the "window-rattling sound", whose volume becomes higher when the user gets closer to the window, as depicted in Figure 13 (e, 2). Moreover, P1 added that even more non-visual mappings are possible by stating "[a] pathway to tactile would be cool" and then described how one could use a haptic device that generates wind to mask any cold temperature. The idea of leveraging multimodal feedback to map distractions is a promising future direction for our technique.

6. Explanation after the fact. We observed participants leveraging "hindsight" in certain designs. In these designs, the moment (cause) that leads to the external stimuli (effect) is unseen by the player and can only be understood after it has happened. For instance, P2 made one of the museum paintings fall after any long sound was heard by the participants, as shown in Figure 13 (b, 2). We also used this technique in Study #1: the participant felt a tap on their back, but they only saw a ninja running away when looked back —they never actually saw the ninja touching but attributed the "tap" to it, only in *hindsight*. While this technique also provides a way to map distractions, it requires specific situations where hindsight can be robustly evoked; as such, it is less generalizable—this technique was used in only 3 out of 36 designs in Study #3.

Qualitative feedback on using our concept. All the participants reacted positively to the concept of integrating distractions into virtual reality. P1, for instance, stated: "[I] really liked the concept and can tell it to be an add-on tool for creating VR experience." Similarly, P3 found the concept useful and stated: "we already got scared when we heard a sudden noise in VR as we can't see it. Having full immersion from inside the game and technically outside the game makes things better." P2 added by describing their design

experience as easy, "most of the time I spent was conceptualizing, once I had an idea, it's pretty easy to get it running."

10 DESIGN CONSIDERATIONS

Based on our studies' results, we synthesized a preliminary set of design considerations. We believe this initial set of design suggestions can serve as a starting point for future VR content creators working toward integrating distractions into virtual reality.

1. Make mappings inconsequential to the VR experience. We encourage prospective designers to make their VR effects (that mask a distraction) inconsequential to the VR experience. This allows the VR experience to be repeatable regardless of the invoked distractions, which can be varied on different "runs" of the same VR experience. For instance, we advise that the VR effects only mask the source of the stimulus but do not provide users with extra information, extra points, extra items, etc. Obviously, our technique can also open some creative opportunities in which designers purposefully leverage the distractions to make the VR experience contextual and unique, e.g., a user would need to experience wind to unlock a special alternate ending, etc.

2. Minimize work using one-to-many mappings. We encourage prospective designers to minimize the amount of VR effects required by designing effects that can be applied to multiple sources of distraction under the same type. We also observed this in our user study with VR designers and recommend adopting this technique. Moreover, this technique is enhanced by the concept of *stretch* mapping, which our study also found to minimize distractions when compared to just noise-canceling headphones. A more extreme way to do this is reusing integration mappings for different types of stimuli.

3. Utilize alternative designs. When possible, we encourage prospective designers to create alternative or parametric VR effects for the same distraction. This allows to keep repetition to a minimum and creates a sense of a lively and responsive VR experience. This suggestion, however, should be balanced with the amount of work (and costs) since it requires designers to create more VR effects for each distraction.

4. Do not overload the user with VR effects. We encourage prospective designers to render a limited set of VR effects at a time. In a realistic setting, such as the outdoor environments in which we conducted Study #2, distractions can occur often, even simultaneously (two distractions being detected almost at the same time) or overlapping one another (a new distraction being detected while another one is still ongoing). While the most straightforward approach is to simply trigger all the VR counterparts of any detected distraction, this might not result in the best experience for the user; in other words, the distraction mappings themselves can become a distraction. One way designers can tackle this is by being mindful to not overload users with too many distraction mappings, for instance: by defining a maximum amount of simultaneous mappings, defining the minimum timing between mappings, etc.

5. Consider impacts on VR narrative before designing mappings. We encourage prospective designers to assess the impact that the mappings will have on VR narrative and evaluate the costs of adjusting the narrative to fit more distractions. We believe that the most common usage will be to align the distractions to the

narrative and not vice-versa, due to costs. As such, we recommend that, in most average cases, prospective designers only include distractions that are sensible to their narrative, rather than engaging in complex narrative changes that are costly.

6. Handle false positives over false negatives. We encourage prospective designers to consider how they handle false positives (i.e., instances where the hardware detected a distraction where there was none in the physical environment). One way to tackle these is to create designs that are very sensible in VR and will not fill out of place even in the absence of external stimuli. Another, more advanced way, is to complement each design with additional multimodal cues in the VR. For instance, in Study #3, one of our designers created a "ceiling fan" to mask wind stimuli. However, their VR ceiling fan also had sound, which makes this design more self-consistent even in the absence of wind. Conversely, we recommend that prospective designers not obsess about false negatives (i.e., instances where the hardware module missed a detection) as this is what VR users already are accustomed to with today's VR (i.e., they know that there are a lot of distractions around them).

7. Ethical considerations. We urge designers to take ethical considerations seriously while integrating distractions in VR. While our proposal ensures the user has full agency in deciding which stimuli they allow to be mapped at the start of each VR experience, we still encourage prospective designers and researchers to consider ethically what is being mapped. For instance, we do not recommend mapping emergency events such as alarms or sirens. Moreover, several distractions carry different values at different ranges. For instance, for a VR user in their home environment, cold temperature shifts are relatively harmless (e.g., A/C or a window was opened, etc.), compared to an expected heat shift (e.g., most likely the A/C) but at a temperature above 35C could be indicative of hazards. Note that while traditional VR also suffers from these safety issues, it is even more important to consider it when integrating distractions. In any implementation, just as in ours, we urge that users are left with a full agency to decide which distractions will be integrated.

11 CONCLUSION

We proposed, implemented, and validated a new concept, where we directly integrate distractions from the user's surroundings into their virtual reality experience to enhance presence. Using our approach, a sensation that contradicts with VR experience (e.g., a drafty wind through the user's room) can be directly mapped to a virtual effect coherent with the VR narrative (e.g., the sway of trees in a VR experience that contains trees).

Using this novel approach, we demonstrated how to integrate a range of distractive stimuli into the VR experience, such as haptics (temperature, vibrations, touch), sounds, and smells. We validated our approach by means of three user studies and a technical evaluation, including a controlled study, an outside-the-lab study, and a study with VR game designers who might one day utilize a system inspired by our findings.

Finally, we synthesized the insights from our studies and explorations into a set of design suggestions and discuss ethical considerations of integrating real-world stimuli into virtual reality.

UIST '22, October 29-November 02, 2022, Bend, OR, USA

ACKNOWLEDGMENTS

Thanks to our colleagues at the Human Computer Integration Lab for participating in numerous pilot studies and for their feedback. Special thanks to Alan Shen, Romain Nith, and Ruoda Yuan for support on videos and figures. This work was supported by NSF grant 2047189. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of any funding agencies.

REFERENCES

- [1] [1] Christopher C. Berger, Mar Gonzalez-Franco, Eyal Ofek, and Ken Hinckley. 2018. The uncanny valley of haptics. *Sci. Robot.* 3, 17 (April 2018), eaar7010. DOI:https://doi.org/10.1126/scirobotics.aar7010
- [2] [2] John M. Carroll, Robert L. Mack, and Wendy A. Kellogg. 1988. Chapter 3 - Interface Metaphors and User Interface Design. In *Handbook of Human-Computer Interaction*, MARTIN Helander (ed.). North-Holland, Amsterdam, 67–85. DOI:https://doi.org/10.1016/B978-0-444-70536-5.50008-7
- [3] [3] Lung-Pan Cheng, Eyal Ofek, Christian Holz, Hrvoje Benko, and Andrew D. Wilson. 2017. Sparse Haptic Proxy: Touch Feedback in Virtual Environments Using a General Passive Prop. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, ACM, Denver Colorado USA, 3718–3728. DOI:https://doi.org/10.1145/3025453.3025753
- [4] [4] Lung-Pan Cheng, Eyal Ofek, Christian Holz, and Andrew D. Wilson. 2019. VRoamer: Generating On-The-Fly VR Experiences While Walking inside Large, Unknown Real-World Building Environments. In 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), IEEE, Osaka, Japan, 359–366. DOI:https: //doi.org/10.1109/VR.2019.8798074
- [5] [5] Isamu Endo, Kazuki Takashima, Maakito Inoue, Kazuyuki Fujita, Kiyoshi Kiyokawa, and Yoshifumi Kitamura. 2021. A Reconfigurable Mobile Head-Mounted Display Supporting Real World Interactions. In Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–7. Retrieved March 30, 2022 from https://doi.org/10.1145/3411765.3451765
- [6] [6] Isamu Endo, Kazuki Takashima, Maakito Inoue, Kazuyuki Fujita, Kiyoshi Kiyokawa, and Yoshifumi Kitamura. 2021. ModularHMD: A Reconfigurable Mobile Head-Mounted Display Enabling Ad-hoc Peripheral Interactions with the Real World. In The 34th Annual ACM Symposium on User Interface Software and Technology, ACM, Virtual Event USA, 100–117. DOI:https://doi.org/10.1145/ 3472749.3474738
- [7] [7] Maia Garau, Doron Friedman, Hila Ritter Widenfeld, Angus Antley, Andrea Brogni, and Mel Slater. 2008. Temporal and Spatial Variations in Presence: Qualitative Analysis of Interviews from an Experiment on Breaks in Presence. *Presence: Teleoperators and Virtual Environments* 17, 3 (June 2008), 293–309. DOI:https://doi.org/10.1162/pres.17.3.293
- [8] [8] Ceenu George, Philipp Janssen, David Heuss, and Florian Alt. 2019. Should I Interrupt or Not? Understanding Interruptions in Head-Mounted Display Settings. In Proceedings of the 2019 on Designing Interactive Systems Conference (DIS '19), Association for Computing Machinery, New York, NY, USA, 497–510. DOI:https: //doi.org/10.1145/3322276.3322363
- [9] [9] Sarthak Ghosh, Lauren Winston, Nishant Panchal, Philippe Kimura-Thollander, Jeff Hotnog, Douglas Cheong, Gabriel Reyes, and Gregory D. Abowd. 2018. NotifiVR: Exploring Interruptions and Notifications in Virtual Reality. *IEEE Trans. Visual. Comput. Graphics* 24, 4 (April 2018), 1447–1456. DOI:https: //doi.org/10.1109/TVCG.2018.2793698
- [10] [10] Matt Gottsacker, Nahal Norouzi, Kangsoo Kim, Gerd Bruder, and Greg Welch. 2021. Diegetic Representations for Seamless Cross-Reality Interruptions. In 2021 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), IEEE, Bari, Italy, 310–319. DOI:https://doi.org/10.1109/ISMAR52148.2021.00047
- [11] [11] Yizheng Gu, Chun Yu, Zhipeng Li, Weiqi Li, Shuchang Xu, Xiaoying Wei, and Yuanchun Shi. 2019. Accurate and Low-Latency Sensing of Touch Contact on Any Surface with Finger-Worn IMU Sensor. In Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, ACM, New Orleans LA USA, 1059–1070. DOI:https://doi.org/10.1145/3332165.3347947
- [12] [12] Daniel Harley, Alexander Verni, Mackenzie Willis, Ashley Ng, Lucas Bozzo, and Ali Mazalek. 2018. Sensory VR: Smelling, Touching, and Eating Virtual Reality. In Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction, ACM, Stockholm Sweden, 386–397. DOI:https://doi.org/10. 1145/3173225.3173241
- [13] [13] Jeremy Hartmann, Christian Holz, Eyal Ofek, and Andrew D. Wilson. 2019. RealityCheck: Blending Virtual Environments with Situated Physical Reality. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, ACM, Glasgow Scotland Uk, 1–12. DOI:https://doi.org/10.1145/3290605.3300577

- [14] [14] Carrie Heeter. 1992. Being There: The Subjective Experience of Presence. Presence: Teleoperators & Virtual Environments 1, 2 (January 1992), 262–271. DOI:https://doi.org/10.1162/pres.1992.1.2.262
- [15] K. Hinckley, R. Pausch, J. Goble, and N. Kassell. 1994. Passive real-world interface props for neurosurgical visualization. In *Proceedings of the SIGCHI* conference on Human factors in computing systems, 452–458. DOI:https://doi.org/ 10.1145/191666.191821
- [16] [16] H.G. Hoffman. 1998. Physically touching virtual objects using tactile augmentation enhances the realism of virtual environments. In *Proceedings. IEEE* 1998 Virtual Reality Annual International Symposium (Cat. No.98CB36180), 59–63. DOI:https://doi.org/10.1109/VRAIS.1998.658423
- [17] [17] Ching-Yu Hsieh, Yi-Shyuan Chiang, Hung-Yu Chiu, and Yung-Ju Chang. 2020. Bridging the Virtual and Real Worlds: A Preliminary Study of Messaging Notifications in Virtual Reality. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–14. Retrieved March 29, 2022 from https://doi.org/10.1145/ 3313831.3376228
- [18] Gierad Laput, Karan Ahuja, Mayank Goel, and Chris Harrison. 2018. Ubicoustics: Plug-and-Play Acoustic Activity Recognition. In Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology, ACM, Berlin Germany, 213–224. DOI:https://doi.org/10.1145/3242587.3242609
- [19] [19] Bernhard Lehner, Gerhard Widmer, and Sebastian Bock. 2015. A low-latency, real-time-capable singing voice detection method with LSTM recurrent neural networks. In 2015 23rd European Signal Processing Conference (EUSIPCO), 21–25. DOI:https://doi.org/10.1109/EUSIPCO.2015.7362337
- [20] Huimin Liu, Zhiquan Wang, Angshuman Mazumdar, and Christos Mousas. 2021. Virtual reality game level layout design for real environment constraints. *Graphics and Visual Computing* 4, (June 2021), 200020. DOI:https://doi.org/10. 1016/j.gvc.2021.200020
- [21] [21] Sebastian Marwecki and Patrick Baudisch. 2018. Scenograph: Fitting Real-Walking VR Experiences into Various Tracking Volumes. In Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology, ACM, Berlin Germany, 511–520. DOI:https://doi.org/10.1145/3242587.3242648
- [22] [22] Brett R. C. Molesworth, Marion Burgess, and Daniel Kwon. 2013. The use of noise cancelling headphones to improve concurrent task performance in a noisy environment. *Applied Acoustics* 74, 1 (January 2013), 110–115. DOI:https: //doi.org/10.1016/j.apacoust.2012.06.015
- [23] [23] Yves Moser and Martin A. M. Gijs. 2007. Miniaturized Flexible Temperature Sensor. Journal of Microelectromechanical Systems 16, 6 (December 2007), 1349– 1354. DOI:https://doi.org/10.1109/JMEMS.2007.908437
- [24] [24] Adiyan Mujibiya, Xiang Cao, Desney S. Tan, Dan Morris, Shwetak N. Patel, and Jun Rekimoto. 2013. The sound of touch: on-body touch and gesture sensing based on transdermal ultrasound propagation. In *Proceedings of the 2013 ACM international conference on Interactive tabletops and surfaces*, ACM, St. Andrews Scotland, United Kingdom, 189–198. DOI:https://doi.org/10.1145/2512349.2512821
 [25] [25] Donald A. Norman. 1999. Affordance, conventions, and design. interactions
- [25] [25] Donald A. Norman. 1999. Affordance, conventions, and design. *Interactions* 6, 3 (May 1999), 38–43. DOI:https://doi.org/10.1145/301153.301168
- [26] [26] Catherine Oh, Fernanda Herrera, and Jeremy Bailenson. 2019. The Effects of Immersion and Real-World Distractions on Virtual Social Interactions. *Cyberpsychology, Behavior, and Social Networking* 22, 6 (June 2019), 365–372. DOI:https://doi.org/10.1089/cyber.2018.0404
- [27] [27] Joseph O'Hagan, Julie R. Williamson, and Mohamed Khamis. 2020. Bystander interruption of VR users. In Proceedings of the 9TH ACM International Symposium on Pervasive Displays, ACM, Manchester United Kingdom, 19–27. DOI:https: //doi.org/10.1145/3393712.3395339
- [28] Sigeru Omatu and Mitsuaki Yano. 2016. E-nose system by using neural networks. *Neurocomputing* 172, (January 2016), 394–398. DOI:https://doi.org/10. 1016/j.neucom.2015.03.101
- [29] [29] Craig Phillips. 2018. Biomimicry in UX. Medium. Retrieved April 5, 2022 from https://medium.com/@craig5446/biomimicry-in-ux-design-249cc6d8649
- [30] [30] Mirco Rossi, Sebastian Feese, Oliver Amft, Nils Braune, Sandro Martis, and Gerhard Tröster. 2013. AmbientSense: A real-time ambient sound recognition system for smartphones. In 2013 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops), 230–235. DOI:https://doi.org/10.1109/PerComW.2013.6529487
- [31] [31] Rufat Rzayev, Sven Mayer, Christian Krauter, and Niels Henze. 2019. Notification in VR: The Effect of Notification Placement, Task and Environment. In Proceedings of the Annual Symposium on Computer-Human Interaction in Play (CHI PLAY '19), Association for Computing Machinery, New York, NY, USA, 199–211. DOI:https://doi.org/10.1145/3311350.3347190
- [32] [32] Munehiko Sato, Ivan Poupyrev, and Chris Harrison. 2012. Touché: enhancing touch interaction on humans, screens, liquids, and everyday objects. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, Austin Texas USA, 483–492. DOI:https://doi.org/10.1145/2207676.2207743
- [33] [33] Lior Shapira and Daniel Freedman. 2016. Reality Skins: Creating Immersive and Tactile Virtual Environments. In 2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), IEEE, Merida, Yucatan, Mexico, 115–124. DOI:https://doi.org/10.1109/ISMAR.2016.23

- [34] [34] Adalberto L. Simeone, Eduardo Velloso, and Hans Gellersen. 2015. 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, ACM, Seoul Republic of Korea, 3307–3316. DOI:https://doi.org/10.1145/2702123.2702389
- [35] [35] Mel Slater. A note on presence terminology. Presence connect 3, 3, 1-5.
- [36] [36] Mel Slater, Andrea Brogni, and Anthony Steed. 2003. Physiological Responses to Breaks in Presence: A Pilot Study. In In The 6th Annual International Workshop on Presence, 3–23.
- [37] [37] Mel Slater and Anthony Steed. 2000. A Virtual Presence Counter. Presence: Teleoperators & Virtual Environments 9, 5 (October 2000), 413–434. DOI:https: //doi.org/10.1162/105474600566925
- [38] [38] Mel Slater and Sylvia Wilbur. 1997. A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. Presence: Teleoperators & Virtual Environments 6, 6 (December 1997), 603–616. DOI:https://doi.org/10.1162/pres.1997.6.6.603
- [39] [39] Misha Sra, Sergio Garrido-Jurado, Chris Schmandt, and Pattie Maes. 2016. Procedurally generated virtual reality from 3D reconstructed physical space. In Proceedings of the 22nd ACM Conference on Virtual Reality Software and Technology, ACM, Munich Germany, 191–200. DOI:https://doi.org/10.1145/2993369.2993372
- [40] [40] Li Su, Hailu Wang, Zhen Tian, Haojie Wang, Qian Cheng, and Wei Yu. 2019. Low Detection Limit and High Sensitivity Wind Speed Sensor Based on Triboelectrification-Induced Electroluminescence. Advanced Science 6, 23 (2019), 1901980. DOI:https://doi.org/10.1002/advs.201901980
- [41] [41] Qi Sun, Li-Yi Wei, and Arie Kaufman. 2016. Mapping virtual and physical reality. ACM Trans. Graph. 35, 4 (July 2016), 1–12. DOI:https://doi.org/10.1145/ 2897824.2925883

- [42] Kien T.P. Tran, Sungchul Jung, Simon Hoermann, and Robert W. Lindeman. 2019. MDI: A Multi-channel Dynamic Immersion Headset for Seamless Switching between Virtual and Real World Activities. In 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), 350–358. DOI:https://doi.org/10.1109/VR.2019. 8798240
- [43] [43] Martin Usoh, Ernest Catena, Sima Arman, and Mel Slater. 2000. Using Presence Questionnaires in Reality. Presence: Teleoperators & Virtual Environments 9, 5 (October 2000), 497–503. DOI:https://doi.org/10.1162/105474600566989
- [44] [44] Eva L. Waterworth and John A. Waterworth. 2001. Focus, Locus, and Sensus: The Three Dimensions of Virtual Experience. *CyberPsychology & Behavior* 4, 2 (April 2001), 203–213. DOI:https://doi.org/10.1089/109493101300117893
- [45] [45] Danli Wu, Yu Cheng, Dehan Luo, Kin-Yeung Wong, Kevin Hung, and Zhijing Yang. 2019. POP-CNN: Predicting Odor's Pleasantness with Convolutional Neural Network. arXiv:1903.07821 [cs, stat] (March 2019). Retrieved April 1, 2022 from http://arxiv.org/abs/1903.07821
- [46] [46] Jackie (Junrui) Yang, Christian Holz, Eyal Ofek, and Andrew D. Wilson. 2019. DreamWalker: Substituting Real-World Walking Experiences with a Virtual Reality. In Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology, ACM, New Orleans LA USA, 1093–1107. DOI:https: //doi.org/10.1145/3332165.3347875
- [47] [47] What is Skeuomorphism? Definition from Techopedia. Techopedia.com. Retrieved April 5, 2022 from http://www.techopedia.com/definition/28955/ skeuomorphism
- [48] [48] Towards Balancing VR Immersion and Bystander Awareness | Proceedings of the ACM on Human-Computer Interaction. Retrieved from https://dl.acm.org/ doi/10.1145/3486950